TravelTide Customer Segmentation for Personalized Reward Program

EXECUTIVE SUMMARY

This project aimed to develop actionable customer segments to implement a personalized rewards program for TravelTide, a digital travel booking startup. Using a dataset of over 5,000 customers, the analysis combined rule-based logic with machine learning techniques (K-Means and DBSCAN) to group customers based on behavior, demographics, and spending.

Project Objectives

- Segment customers to enable targeted marketing and loyalty strategies
- Compare segmentation approaches for both business relevance and technical quality

Methods Used

- Rule-Based Segmentation: Business-driven personas
- K-Means Clustering: Machine learning-based segmentation using behavioral variables
- DBSCAN Clustering: Density-based clustering focused on identifying outliers

Key Findings

- K-Means provided the most balanced and actionable segmentation, with six distinct clusters. Personas like "Frequent Business Travelers" and "Luxury Explorers" emerged as valuable targets.
- DBSCAN yielded technically strong but impractical results, clustering 99% of users into one group while isolating just 49 as distinct—useful only for identifying VIPs or anomalies.
- Rule-based segments aligned with some K-Means groups, but lacked nuance in spending and cancellation behavior.

Final Recommendations

- Adopt K-Means segmentation for personalization and loyalty campaigns.
- Prioritize Frequent Business Travelers for premium perks and retention.
- Address churn in the High Cancellation Risk group.
- Use DBSCAN only to monitor VIPs or potential issues based on behavioral outliers.
- Plan next steps: validate segments with additional data, test marketing strategies, and monitor impact.

TravelTide Customer Segmentation for Personalized Reward Program

DETAILED REPORT

1. Introduction & Objectives

TravelTide, a digital travel platform, aims to personalize customer experience through a rewards program. This project analyzed user-level booking data (e.g., flights, sessions, cancellations, spending) to uncover meaningful traveler segments.

2. Methodology

2.1 Rule-Based Segmentation

We created eight initial customer personas using filters on age, marital status, parental status, and average checked bags. This approach resulted in logical segments like:

- **Business Travelers** (not married, frequent flyers)
- **Senior Couples** (older, married, lower activity)
- **Young Adventurers** (single, active but low spend)

However, this approach lacked behavioral nuance, and an overlap between segments limited personalization potential.
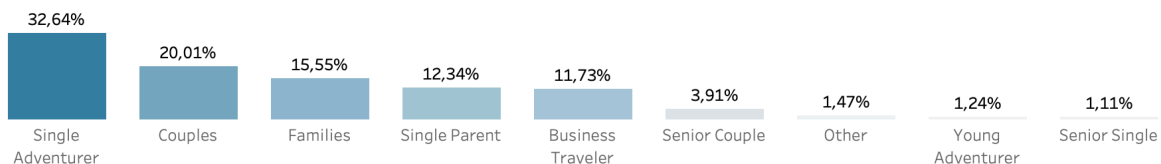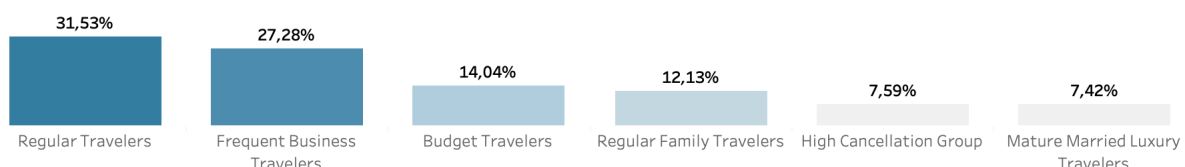
2.2 K-Means Clustering
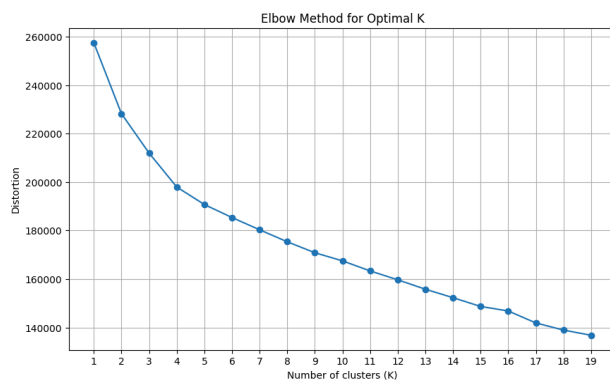
After feature engineering, PCA was used for dimensionality reduction.

The **elbow method** and silhouette score were used to find the optimal number of clusters.

These clusters provided balance (7–31% each) and clear marketing use cases.

TravelTide Customer Segmentation for Personalized Reward Program

The Steady Explorers
Regular Traveler

"I love planning spontaneous getaways — not too fancy, just new experiences and a change of scenery."

✅ Offer targeted destination packages

The Spontaneous Savers
Budget Traveler

"I travel when the deals are good and the vibe feels right — I'm always up for an adventure, as long as it's cheap!"

✅ Offer exclusive discounts or Last-minute deal alerts

The Executive Jetsetters
Business Travelers

"Efficiency, comfort, and staying connected — travel is part of the job, but I expect it to run smoothly."

✅ Offer airport lounge access or Express check-in/security

The Family Explorers
Regular Family Travelers

"Traveling with the kids isn't just a holiday — it's how we make memories that last a lifetime."

✅ Offer free checked bags (they have highest average checked bags)

The Golden Indulgers
Mature Married Luxury Travelers

"We prefer fewer trips, but each one is special — comfort, fine dining, and time to unwind together."

✅ Offer a free hotel night on longer stays (they already stay longest)

The Uncommitted Clickers
High Cancellation Group

"Life changes fast — I book when I can, but I need the flexibility to change my plans without stress."

✅ Offer free cancellation

Six clusters were chosen, each cluster represented a meaningful group, such as:
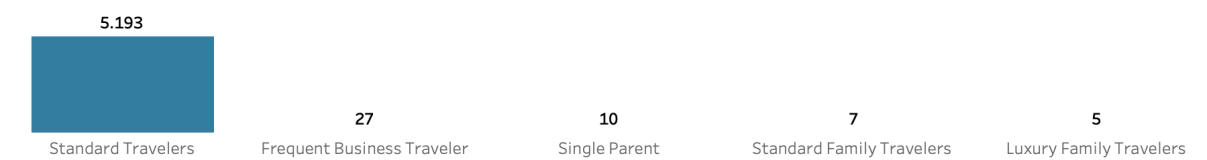
Cluster 1 – Business Travelers: High trips and spend, low cancellation

Cluster 4 – Luxury Explorers: High spenders, many checked bags

Cluster 5 – High Cancellation Risk: Low engagement, high cancellation rate

2.3 DBSCAN Clustering

DBSCAN offered a contrasting approach. It successfully identified **49 clear outliers**, but clustered 99.1% of users into one main group, making it unsuitable for broad segmentation.

5.193	27	10	7	5
Standard Travelers	Frequent Business Traveler	Single Parent	Standard Family Travelers	Luxury Family Travelers

3. Results & Comparison

Method	Segments	Segment Balance	Business Value	Statistical Quality	Notable Value
K-Means	6	🟢Balanced (7–31%)	🟢High	🔴Low (Silhouette ~ 0.1)	🟢Actionable Personas
DBSCAN	5	🔴Skewed (99% in one cluster)	🔴Low	🟢Good (Silhouette ~0.4)	🟡Useful for VIP and outliers
Rule Based	8	🟡Possible overlap	🟡Medium	⚪None	🟡Good starting personas

TravelTide Customer Segmentation for Personalized Reward Program

4. Key Insights

- K–Means segmentation offers the best balance between **actionability and interpretability**.
- DBSCAN, while technically strong, is only useful for **detecting anomalies or VIPs**.
- Rule–based personas are helpful for business communication, but are insufficient for tailored marketing.

5. Recommendations & Next Steps

5.1 Recommendations

- **Adopt K–Means segmentation** to drive the rewards program and marketing strategies.
- **Prioritize Business Travelers** – high revenue, low churn.
- **Develop targeted offers for Luxury Explorers** to boost loyalty.
- **Mitigate churn in the High Cancellation Risk** with retention efforts (e.g., flexible bookings, concierge support).
- **Use DBSCAN insights to design personalized experiences** for identified outliers.

Key Customer Segments		
Business Travelers	**Luxury Explorers**	**High Cancellation Risk**
27.3% of users	8.4% of users	7.6% of users
High revenue, loyal users	High spending users	Frequent cancellation

5.2 Disclaimers & Next Steps

- These clusters are **initial groupings** based on available data.
- Further **validation** is needed with more variables.
- Consider **A/B testing** marketing strategies across segments to measure impact.
- Update segments **periodically** to adapt to shifting customer behavior.

SQL: Cleaned Cohort Selection and Feature Engineering
Data: Cleaned Cohort, Preprocessed Data, Data with Clustering Groups
Notebooks: Session EDA, Cohort EDA, Preprocessing, Clustering, Segments, and Perks