📘 Detailed Report (Expanded Overview)

🎯 Objective

Segment travel users to enable tailored marketing offers that improve conversion, engagement, and retention. Business questions addressed:

- Who are our most valuable users?
- Who is influenced by discounts?
- Who needs re-engagement strategies?

🧱 Methodology

🔍 Data Preparation (SQL)

- Aggregated user activity (sessions, bookings, checked bags, spend, etc.)
- Created views for:
    - Session engagement
    - Travel intensity
    - Discount usage
    - Final merged user profile

🔬 Clustering Approach

- Algorithm: **K-Means**, 3 clusters chosen using business interpretability and silhouette analysis (score = 0.29)
- Features normalized with log scaling and outlier filtering (IQR)
- DBSCAN tested but discarded due to poor grouping

🧩 Segment Profiles

🔵 Cluster 0 – Moderate Travelers

- Moderate engagement (~31 sessions)
- ~16 bookings, ~$17.7K spend
- 36% with children, some discount use
- **Offer**: *Free Checked Bag*

🟢 Cluster 1 – Heavy Spenders & Frequent Flyers

- High engagement: ~81 sessions, ~64 bookings
- ~$65K spend, low discount sensitivity
- Long session durations, high loyalty
- **Offer**: *1-Night Free Hotel*

🟡 Cluster 2 – Light Users & Budget-Conscious

- Low sessions (21), low spend ($7K)
- High price sensitivity, 37% have children
- **Offer**: *Exclusive Discounts or No Cancellation Fees*

📊 Feature Insights

- **Right-skewed spending and session time**: Requires normalization
- **High outliers in bags, spend, and discounts**: Affects segmentation
- **Imbalanced features**: Like cancellations and discount usage

- **Engagement spread**: Clicks and session time vary widely among users

---

📈 Business Recommendations

- **Cluster 1**: Prioritize with loyalty rewards to retain top users
- **Cluster 0**: Upsell with practical benefits to boost spend
- **Cluster 2**: Incentivize re-engagement with price-sensitive offers

---

📂 Next Steps

- Deploy **interactive dashboard** (Streamlit or Tableau)
- Build **user-level recommender** based on cluster score
- Automate updates via **Airflow pipelines**
