



**Data-Driven Personalized Perks to  
Improve Retention**

# **Customer Segmentation**

**WITH THE STUDY PURPOSES**

**Created by Dr. Svitlana Kovalivska**

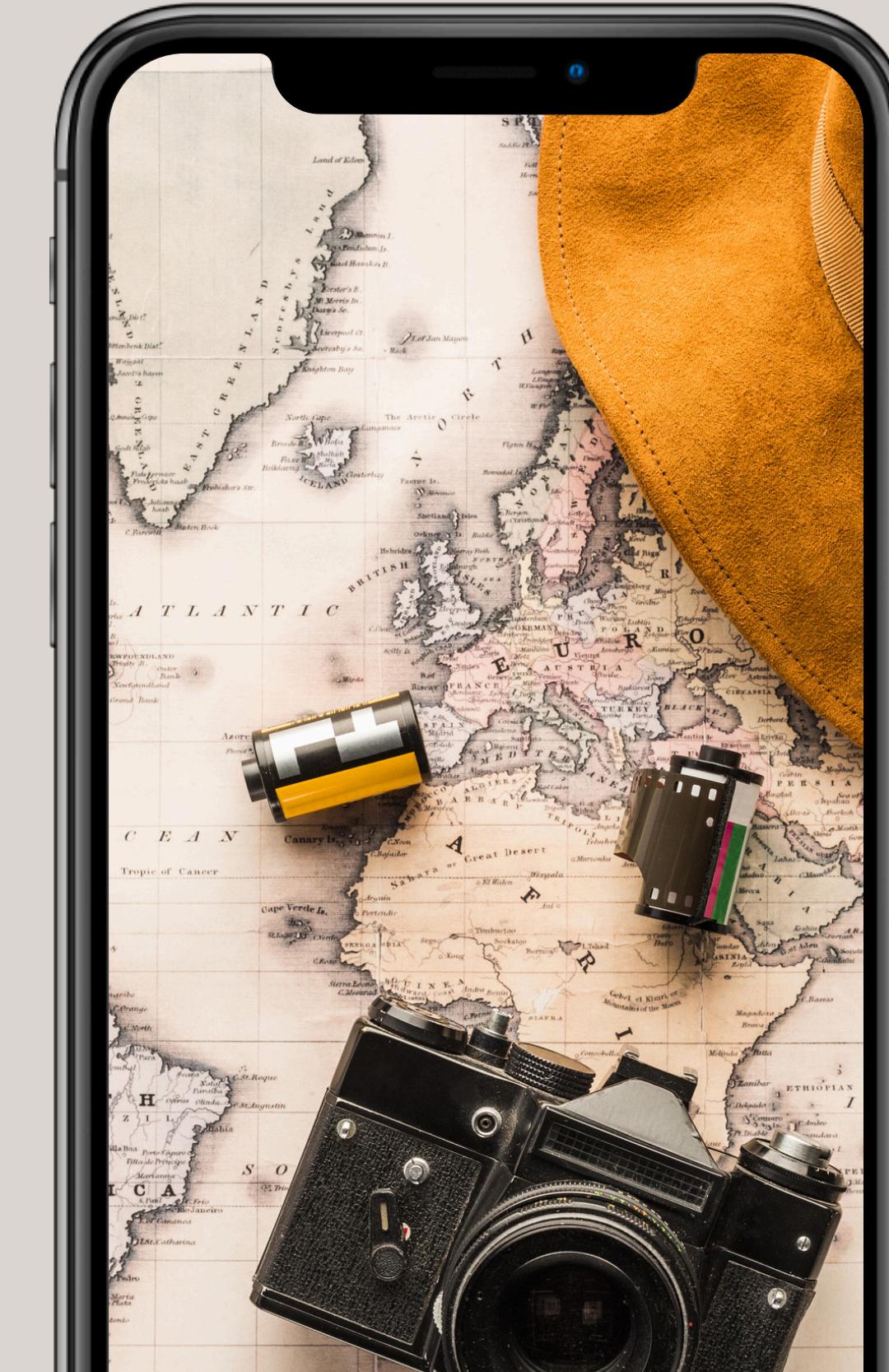


# TravelTide's Business Landscape

Brief Overview: TravelTide is a leading e-booking platform launched in 2021, known for its best-in-class search and data aggregation.

Challenge: Despite a competitive advantage in inventory and technology, customer retention remains a challenge.

Goal: Implement a data-driven personalized rewards strategy to boost repeat bookings.





# Overview

- 01** Introduction
- 02** Problem
- 03** Strategic Objectives
- 04** Data Extraction and Preparation

- 05** Feature Engineering & Data Enhancements
- 06** Data Cleaning & Handling Outliers
- 07** Exploratory Data Analysis Insights
- 08** Data-Driven Customer Segmentation

- 09** Persona Portret
- 10** Perk Distribution & Impact
- 11** Recommendation
- 12** Conclusion

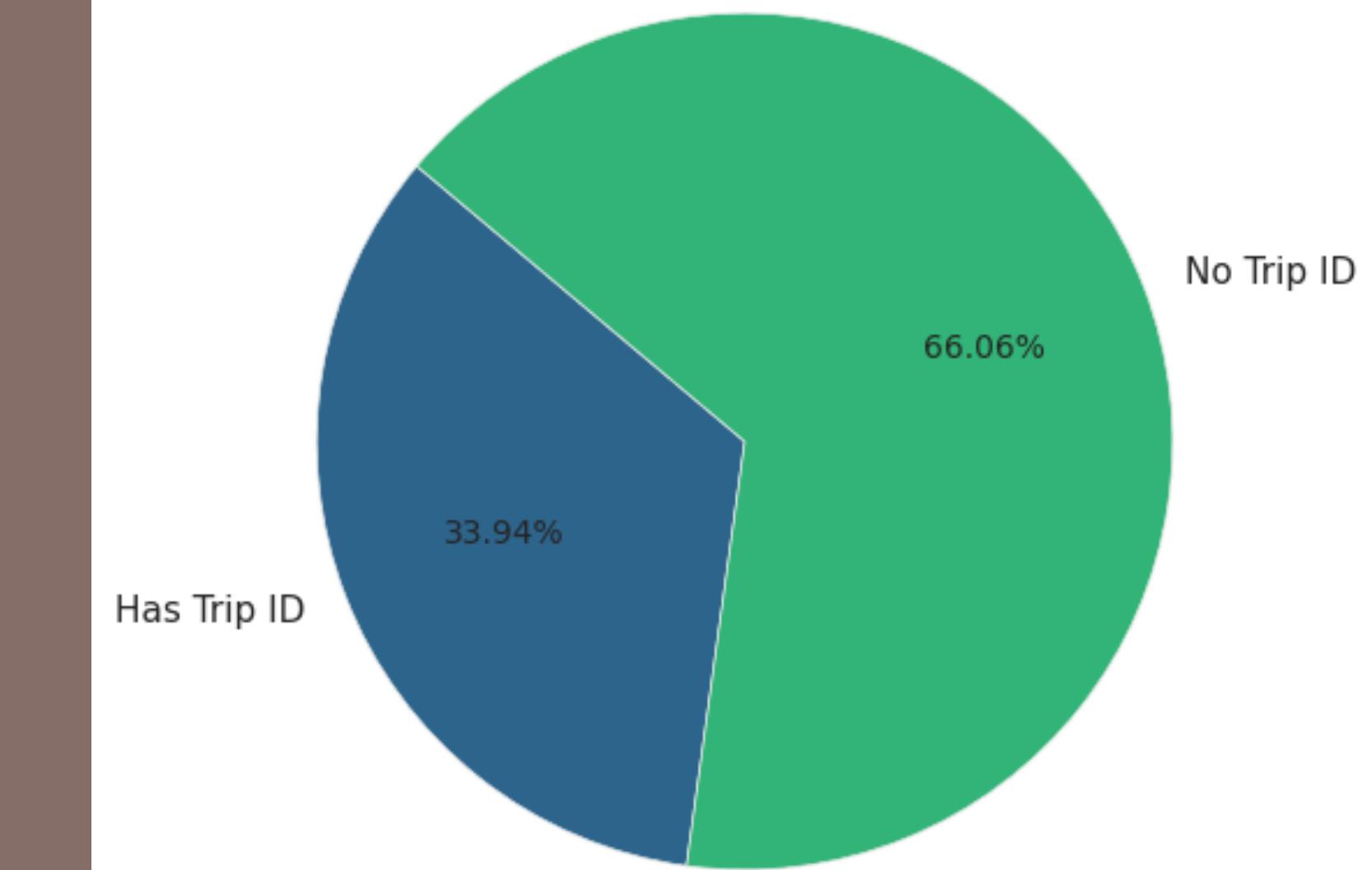
Presented by Dr. Svitlana Kovalivska



# Problem

TravelTide, a fast-growing e-booking platform launched in 2021, excels in travel search and data aggregation but faces challenges in customer retention. A data-driven rewards strategy that personalizes incentives based on customer behaviour was developed to strengthen loyalty and engagement.

Booked vs. Non-booked sessions





# Strategic Objectives

Enhance Retention: Offer targeted, high-value perks based on travel behavior.

Improve Conversion Rates: Replace generic rewards with personalized offers.

Create a Competitive Edge: Leverage customer insights to increase engagement.

## Customer Retention





# Data Extraction and Preparation

## Data Sources:

Travel sessions,  
booking history, and  
customer profiles from  
PostgreSQL.

## Preprocessing:

Data cleaning, missing  
value imputation,  
outlier detection.

## Feature Engineering:

Travel patterns,  
behavioral signals,  
spending habits.



# Feature Engineering & Data Enhancements

## Overview

Introduced more than 20 New Features to improve customer profiling and perk assignment.

### Key Features Added:

- **Fare per Seat:** More accurate relative ticket pricing instead of Base Fare.
- **Lead Time to Departure:** Measures how far in advance bookings are made.
- **Booking Frequency:** Tracks number of bookings within the last 6/12 months.
- **Average Trip Duration:** Helps differentiate between short-term vs. long-haul travelers.
- **Flight-to-Hotel Ratio:** Identifies whether a user books flights alone or as packages.
- **Cancellation Rate:** Determines risk of cancellations for flexible booking perks.
- **Engagement Score:** Based on session duration, clicks, and search activity.



# Data Cleaning & Handling Outliers

01

## Outlier Detection:

- Identified extreme booking behaviors, such as unusually high spending or frequent cancellations.
- Used Interquartile Range (IQR) and Z-score methods to flag outliers.
- Special segmentation applied for high-value travelers (luxury segment).

02

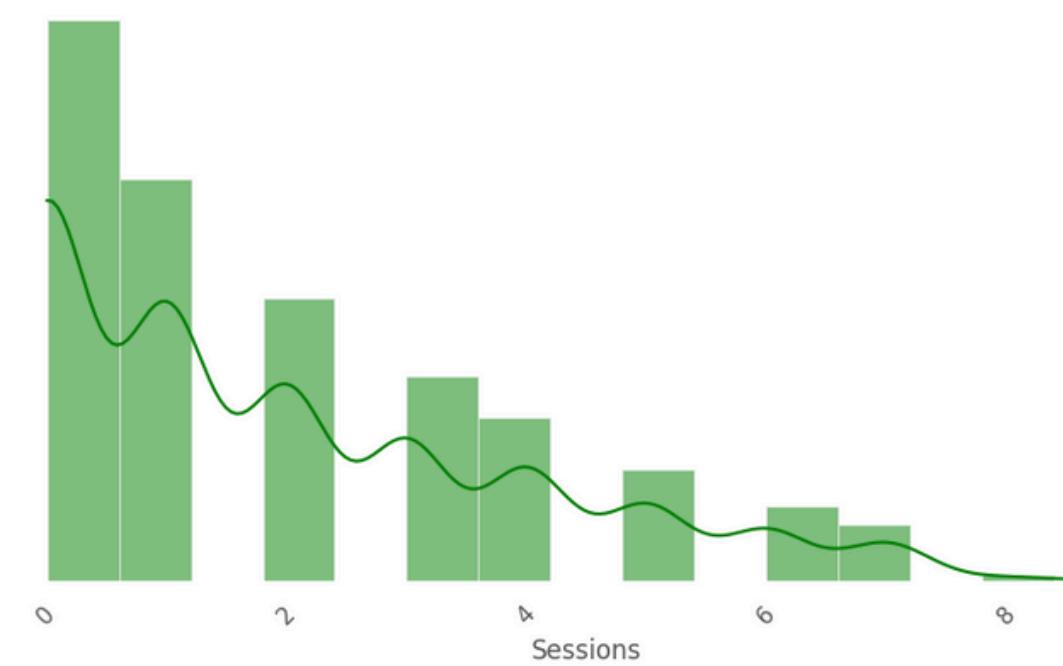
## Handling Negative Night Stays:

- Detected anomalies in check-in and check-out dates (e.g., negative nights due to reversed timestamps).
- Corrected by swapping incorrect values or removing invalid records.
- Ensured clean and reliable data before clustering.

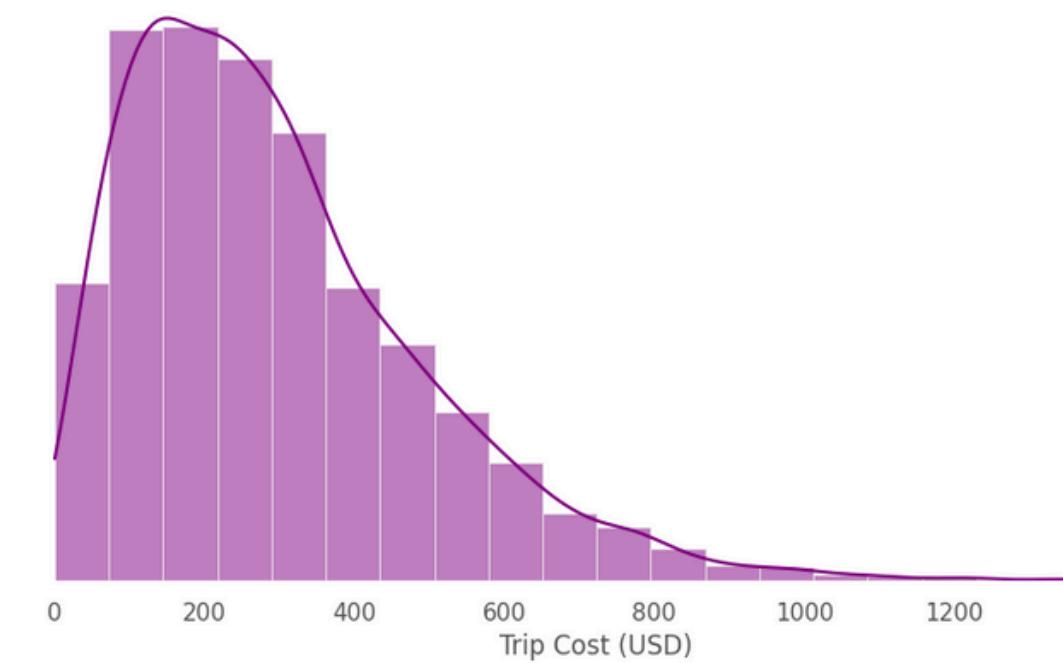


# Exploratory Data Analysis Insights

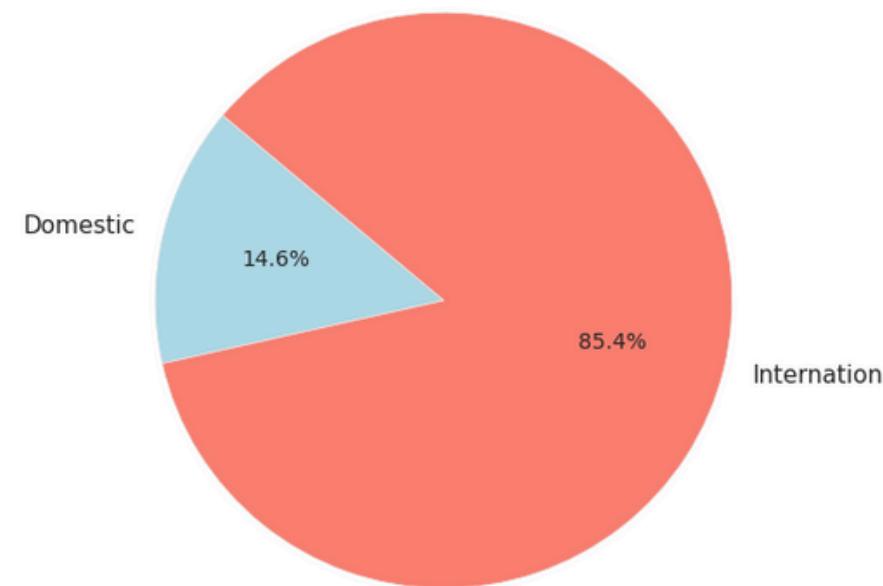
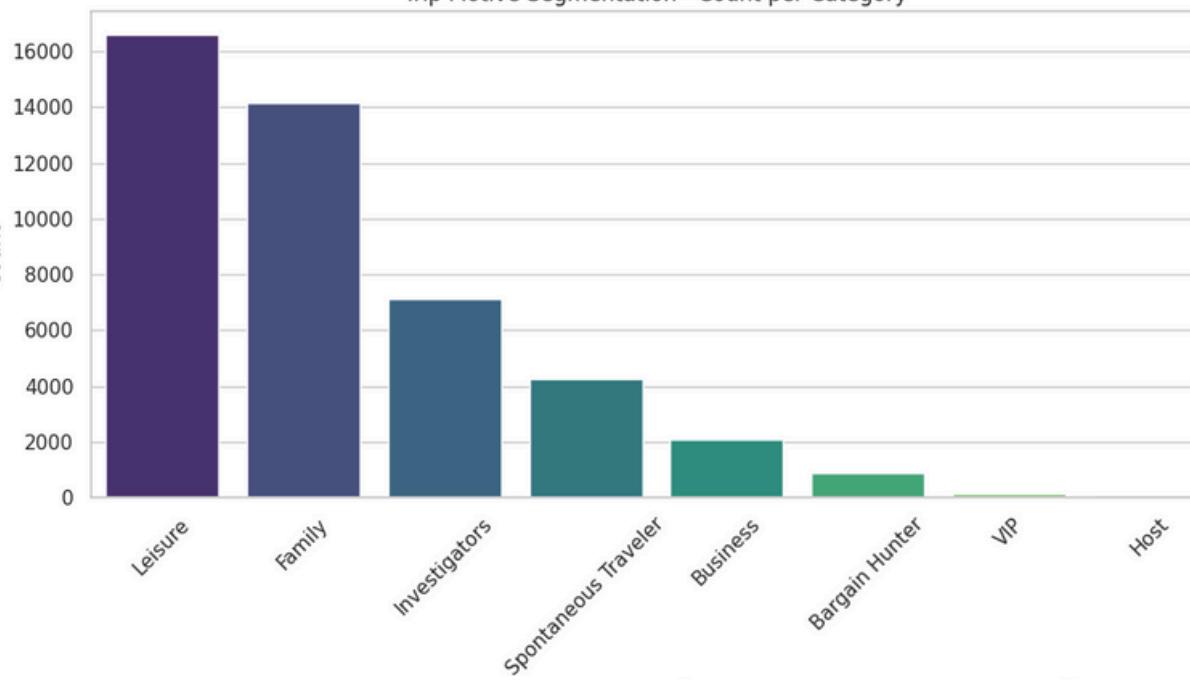
Average Sessions Before First Purchase for Users with Bookings



Distribution of Trip Cost After Discount for Users with Bookings



Trip Motive Segmentation - Count per Category



- Booking Trends: Majority of users browse but don't convert.
- Engagement Metrics: Click rate, session duration, and lead time analyzed.
- Customer Behavior Groups: Identified trends in cost sensitivity, flexibility, and travel frequency.



# Data-Driven Customer Segmentation

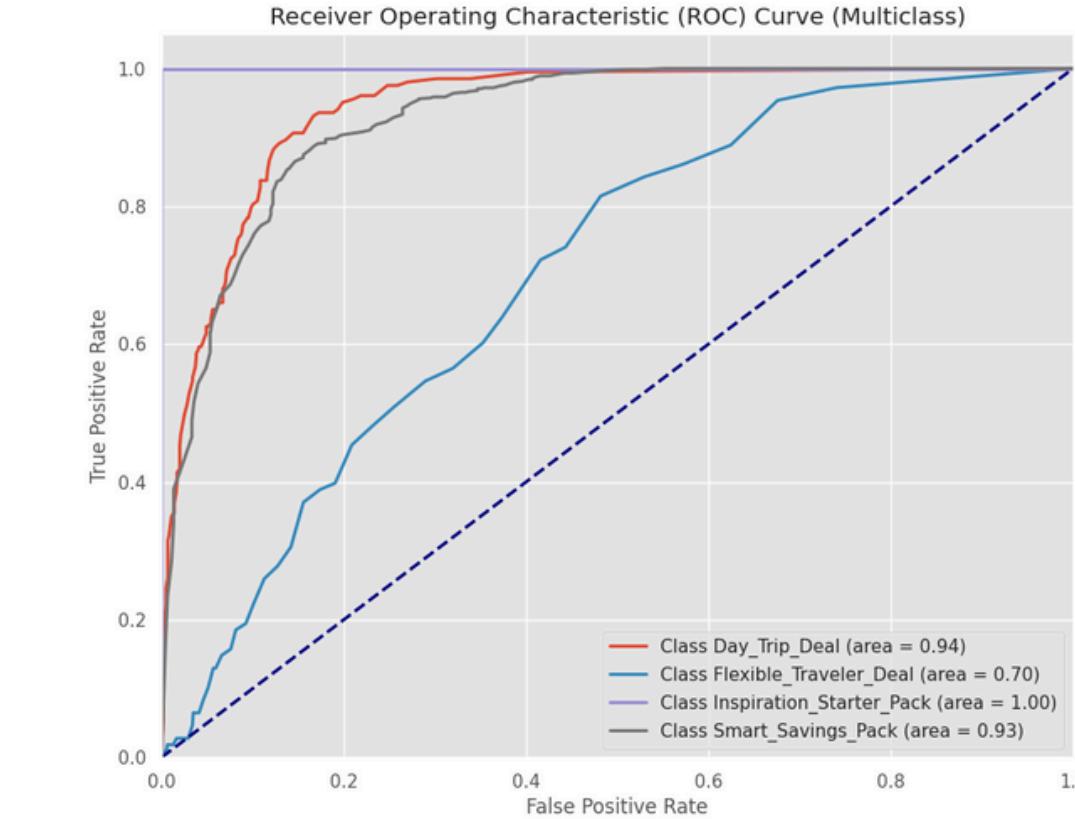
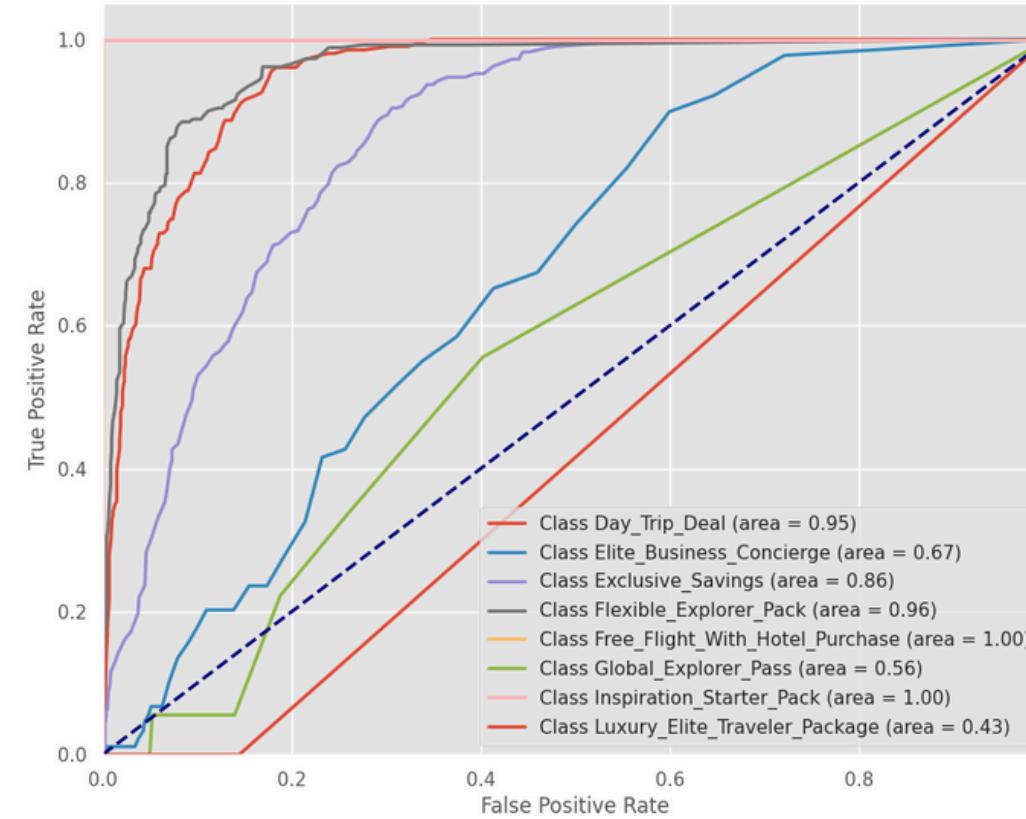
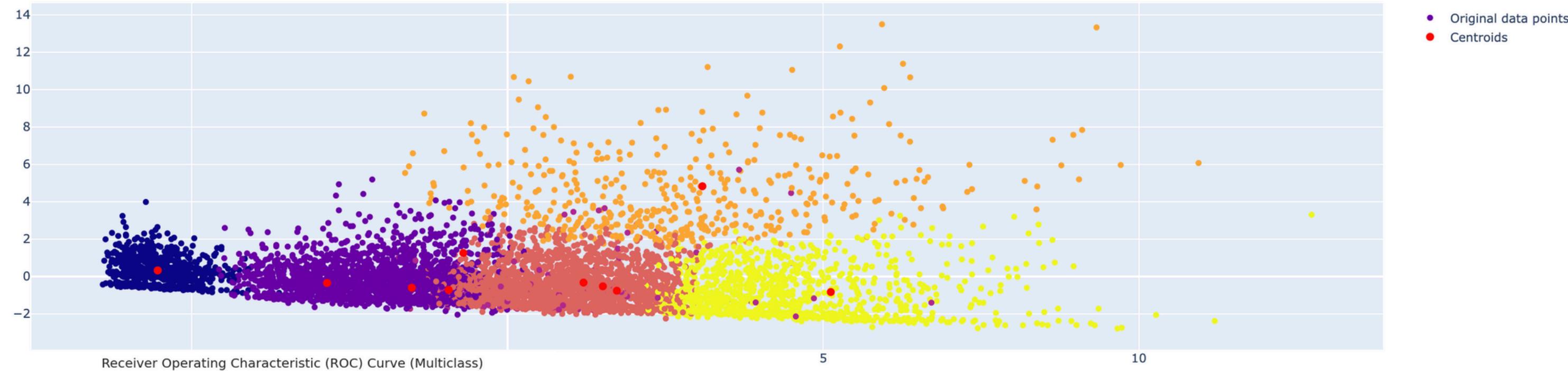
- KMeans clustering identified six traveler segments, later consolidated into four primary groups:



- Methodology, Technologies Used:
- Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn) for data processing, clustering, and visualization.
- SQL (PostgreSQL) for data extraction and transformation.
- Machine Learning (KMeans clustering, PCA for dimensionality reduction) for customer segmentation.



# Clusters





# Persona Portret

## Business Traveler

- User 101961
- Gender and Age: Female, 42 years old.
- User Since: 2022.
- Travel History: 10 flights and 5 hotel stays via TravelTide.
- Travel Characteristics:  
Short stays: average hotel stay duration is 3 nights.  
Minimal luggage: 25% flights with carry-on only.
- Behavior: Prioritizes efficiency and minimizing travel time.
- **Recommended Perk:** VIP Lounge & Priority Check-In.

## Bargain and Price-Sensitive Traveler

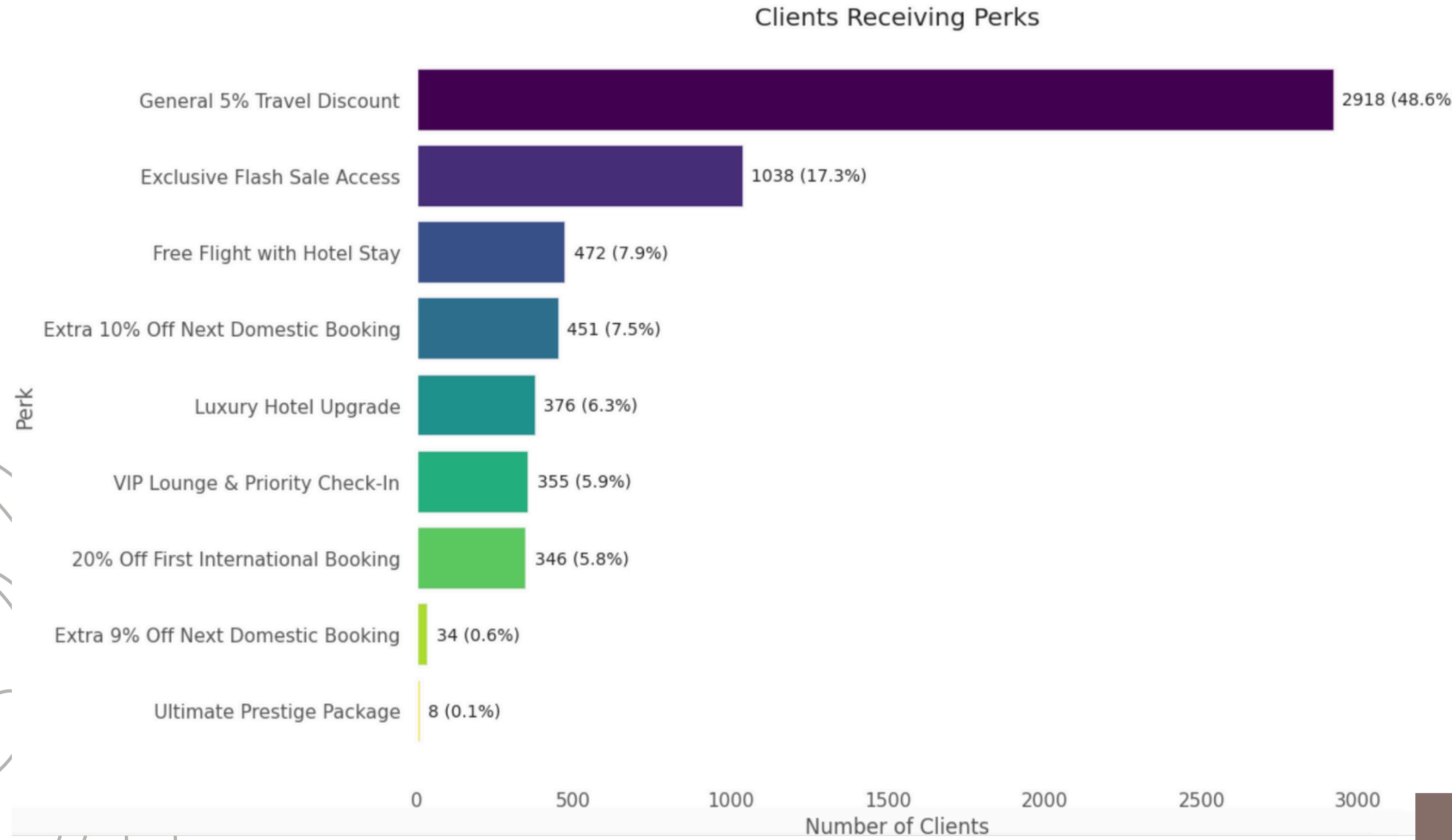
- User 844489
- Gender and Age: Male, 23 years old.
- User Since: 2023.
- Travel History: No trips booked through TravelTide.
- Characteristics:  
Highly interested in deals and discounts.  
Prefers budget hotels and economy-class flights.  
Flexible with travel dates to secure better prices.
- Behavior: Extremely price-sensitive, responds well to discounts.
- **Recommended Perk:** 20% Off First International Booking.

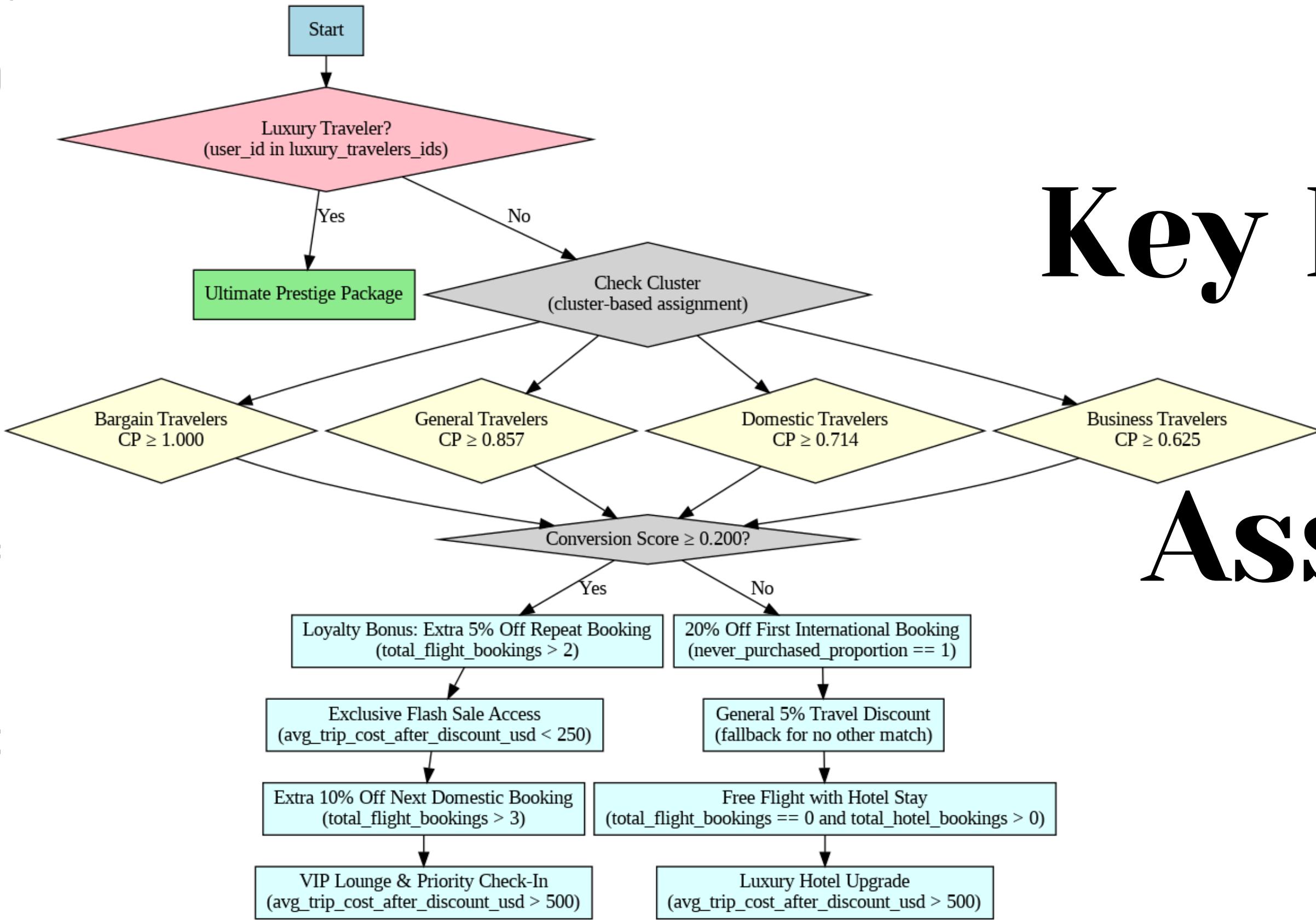
## General Traveler

- User 174997
- Gender and Age: Female, 31 years old.
- User Since: 2020.
- Travel History: 4 flights and 2 hotel stays via TravelTide.
- Travel Characteristics:  
Travels for leisure, including international trips.  
Values flexibility: often books refundable tickets.  
Books both hotels and flights together.
- Behavior: Enjoys traveling but prefers flexibility when booking. Moderate spending level.
- **Recommended Perk:** General 5% Travel Discount.



# Perk Distribution & Impact





# Key Results & Perk Assignment Insights



# Recommendations & Next Steps

## Testing & Optimization Strategy

- A/B Testing: Comparing personalized emails vs. generic rewards emails.
- Key Metrics Monitored: Click-through rates, conversion rates, and booking frequency.
- Expected Outcome: Higher engagement and increased loyalty with personalized perks.

## Implementation Strategy

- Launch Personalized Perk-Based Email Campaigns.
- Conduct A/B Testing on Reward Effectiveness.
- Monitor and Continuously Optimize Perk Assignments.



# Conclusion & Business Impact

## Key Takeaways:

01

Data-driven personalization improves customer retention.

02

Segmentation allows for tailored marketing efforts.

03

Perk optimization leads to increased customer loyalty and conversion.

Next Steps: Implement, measure, and optimize the rewards strategy.



Presented by Dr. Svitlana Kovalivska

---

# Thank You

---

Personalized rewards are the future of TravelTide's customer retention strategy