TravelTide - Customer Segmentation

Date: Sep 25, 2023,
Prepared by: Data Analyst Team

Introduction

This report outlines the findings of the Data Analyst Team's customer segmentation analysis to support implementing a personalized rewards program for TravelTide customers.

TravelTide's commitment to providing an extensive and easily searchable travel inventory has yielded significant benefits; however, certain aspects of the customer experience have room for improvement, impacting customer retention. The Marketing team aims to establish a data-driven rewards program to enhance customer retention and value. This initiative necessitates a comprehensive understanding of customer behavior through effective customer segmentation.

Objectives

The primary objective of the Marketing team is to design an enticing personalized rewards program that resonates with customers, thereby driving platform engagement and customer loyalty.

Goals:

1) Analyze data to validate the hypothesis of customer segments showing a distinct interest in proposed perks.
2) Attribute each customer to a preferred perk category.

Methodology

Customer segmentation divides the customer base into distinct groups, enabling tailored experiences. Our approach leveraged the following methodology:

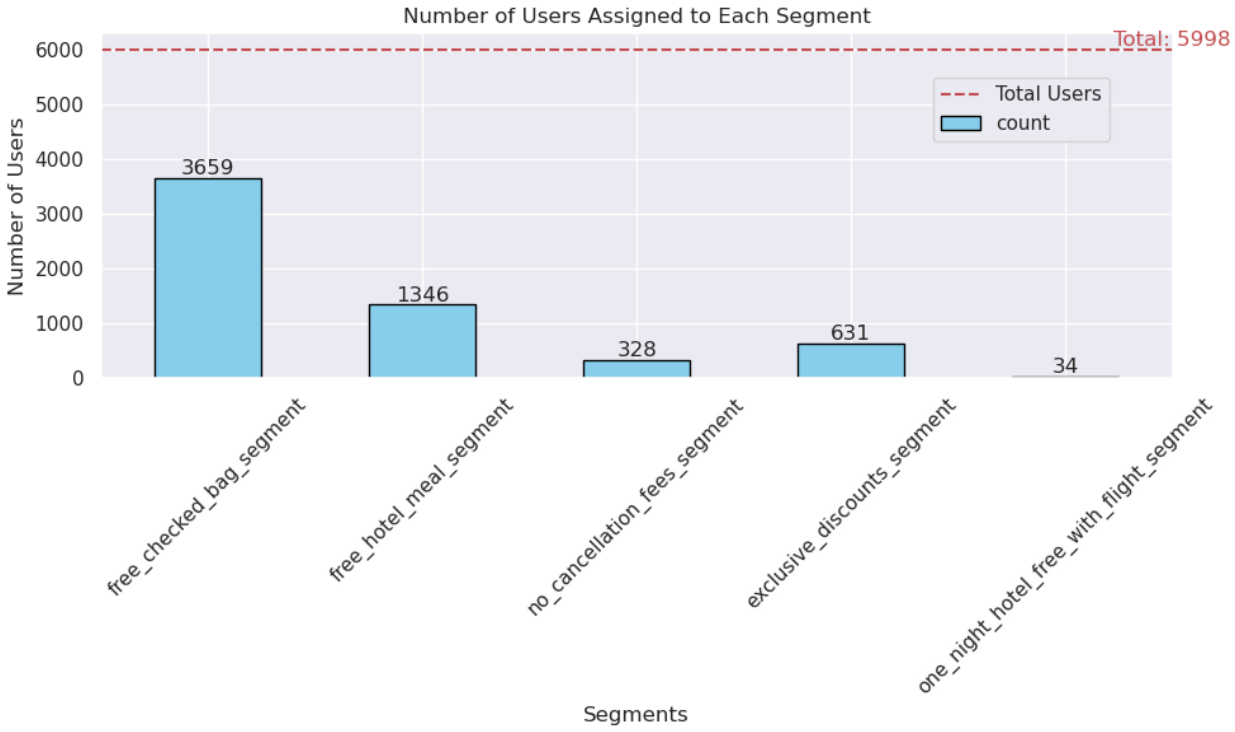
1) **Cohort Definition**: Users engaging in over 7 sessions from '2023-01-04' onwards
2) **Data Extraction**: Employed SQL to extract aggregated cohort data from the PostgreSQL RDBMS system.
3) **Data Analysis**: Utilized SQL and Python to define and rank customer segments.
4) **Segment Assignment**: Assigned each customer to an appropriate segment.

Key Findings

Utilizing the Categorization, Thresholding, and Fuzzy Segmentation techniques, customer segments hypothesized by Marketing were validated and customers were mapped to segments based on their preferences for proposed perks:

- Free Checked Bag Segment: **3,659** customers
- Free Hotel Meal Segment: **1,346** customers
- No Cancellation Fees Segment: **328** customers
- Exclusive Discounts Segment: **631** customers
- One-night Free Hotel with Flight Segment: **34** customers

Note: Users not meeting segment criteria are additionally categorized under the lowest cost segment, 'Free Checked Bag Segment'.



Recommendations and Next Steps

1. **Personalized Marketing Plans**: Craft tailored marketing strategies for each segment utilizing the 'customer_segmented.csv' file. Highlight the perk most appealing to the respective segment.
2. **Cost Consideration**: Assess the feasibility of phased promotions based on the total cost of proposed perks for each segment.

For further details and in-depth analysis, please refer to the accompanying documentation.

Appendix

Appendix 1-Segmentation Criteria and Methodology

Customer Segmentation methods used:

1) **Fuzzy Segmentation** using Categorization and Thresholding
2) **K-means Segmentation -** used for comparing with Fuzzy segmentation

In Fuzzy segmentation, we defined segments in a top-down manner. And then created a ranking order for the segments. Customers were segmented using that ranking order

In the language of formal logic, these methods let us use deductive reasoning. Since we can use deductive reasoning, we can use the knowledge we have about the data to define the segments manually. If we cannot explain used metrics and segments to non-technical audiences in simple words, we will not use those.

K Means belongs to a general class of Machine Learning algorithms called Unsupervised Learners. K Means is a clustering algorithm, which means it groups data into clusters. K Means is a centroid-based algorithm, which means it attempts to minimize the distance between data points and their centroid. It does this by iteratively calculating the centroid of each cluster and reassigning data points to the nearest centroid.

In other words, K-means is a distance-based segmentation that segments customers by behavioral data based on “similarity”, as defined by a distance metric like the Euclidean. Under this perspective, two customers are said to be similar in real life if they are nearby in behavioral data space.

We will go with Fuzzy segmentation as it lets us use deductive reasoning.

Segments definition:

1) **Free checked bag** - If the average number of checked bags is 0 for a user's flights, they will be categorized into the "Free checked bag" segment.
2) **Free hotel meal** - If the average hotel discount amount is 0 for a user, they will be categorized into the "Free hotel meal" segment.
3) **No cancellation fees** - "No cancellation fees" segment: Users with a total sum of cancellations (sessions.cancellation::int) is greater than 0.
4) **Exclusive discounts** - when bargain_hunter_index > 0.0001. (bargain_hunter_index = ASD_per_km*discount_flight_proportion*average_flight_discount)
5) **1-night free hotel with flight** - Users who have booked at least one flight (sessions.flight_booked) and have not booked hotels (sessions.hotel_booked)

Customer segmentation process:

1. Create a ranking order for the segments
2. Segment customers using ranking order

```
# Define the ranking order (Low to High)
RANKING_ORDER = [
    'free_checked_bag_segment',
    'free_hotel_meal_segment',
    'no_cancellation_fees_segment',
    'exclusive_discounts_segment',
    'one_night_hotel_free_with_flight_segment'
]
```

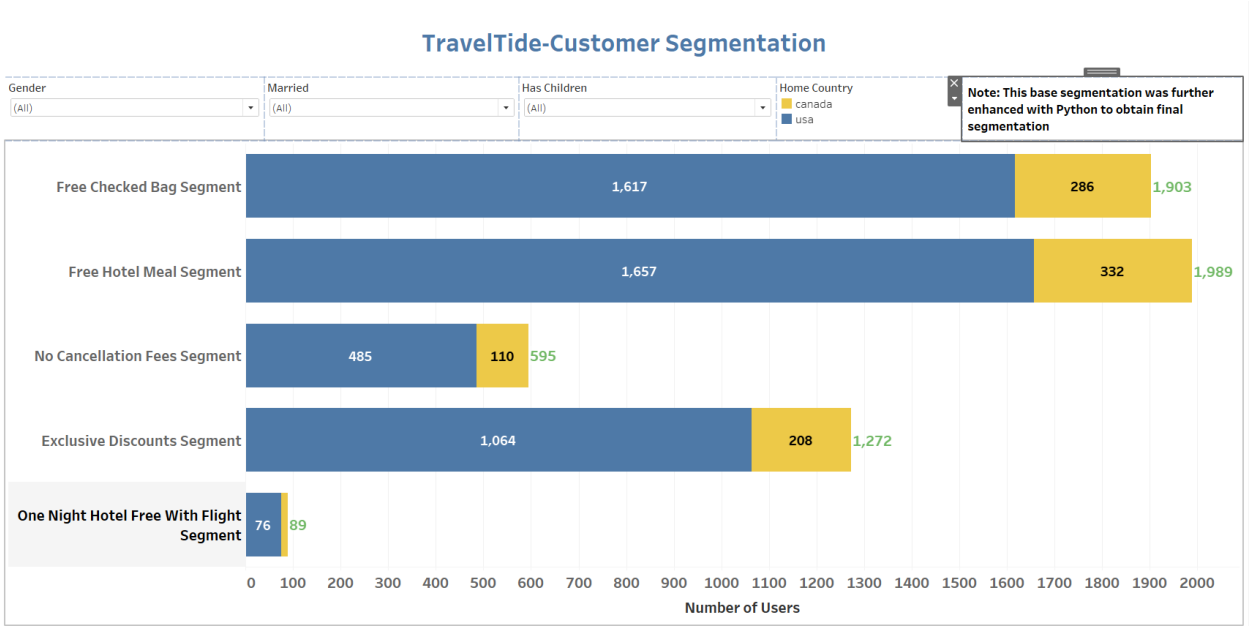
Breakdown of the ranking order:

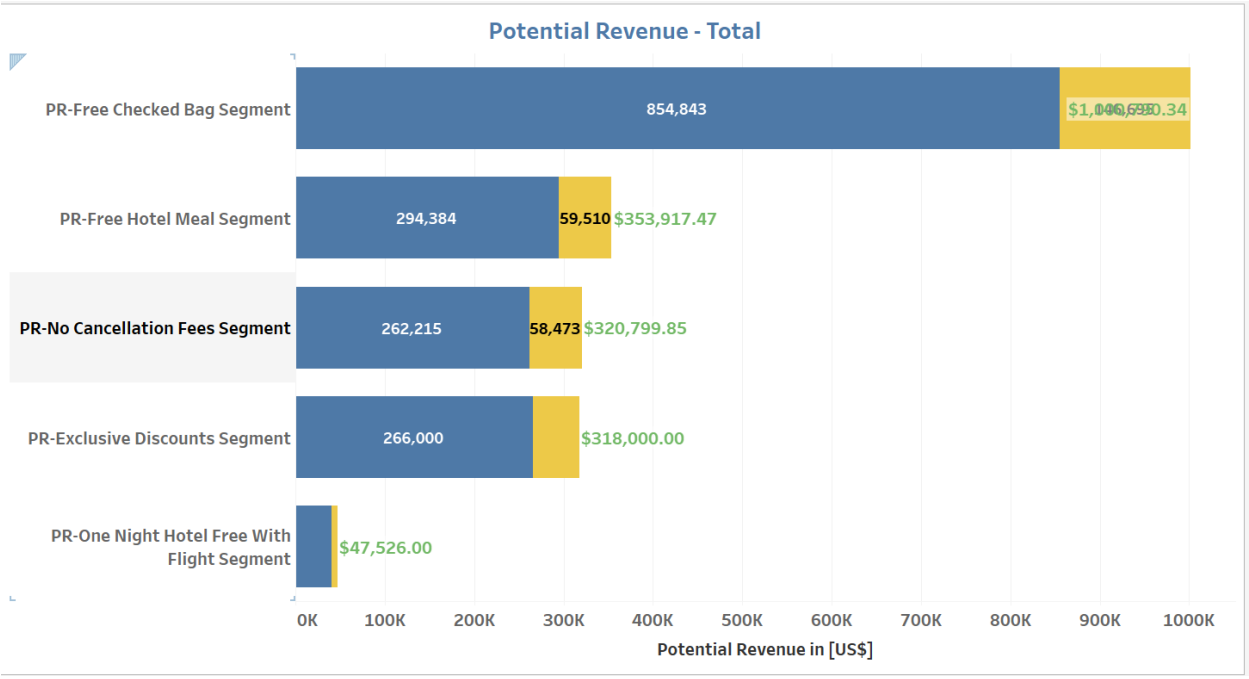
- **Free_checked_bag_segment**: This segment offers a free checked bag, which is a relatively lower-cost benefit than other options. It's a valuable perk for travelers and can attract budget-conscious customers.
- **Free_hotel_meal_segment**: Offering a free hotel meal is also a moderately priced benefit. It can appeal to customers looking to save on dining expenses during their trip.
- **No_cancellation_fees_segment**: This segment promises no cancellation fees. While it might not directly correlate with a specific cost, it can be seen as a customer-friendly policy that adds value, especially for those who value flexibility in their travel plans.
- **Exclusive_discounts_segment**: Exclusive discounts can range in terms of savings, but placing it in the middle of the order suggests it's a more substantial offer than the previous segments.
- **One_night_hotel_free_with_flight_segment**: This segment offers a free hotel night with a flight booking, which can be a higher-cost benefit. It's placed last in the order, implying it's the most valuable or expensive benefit.

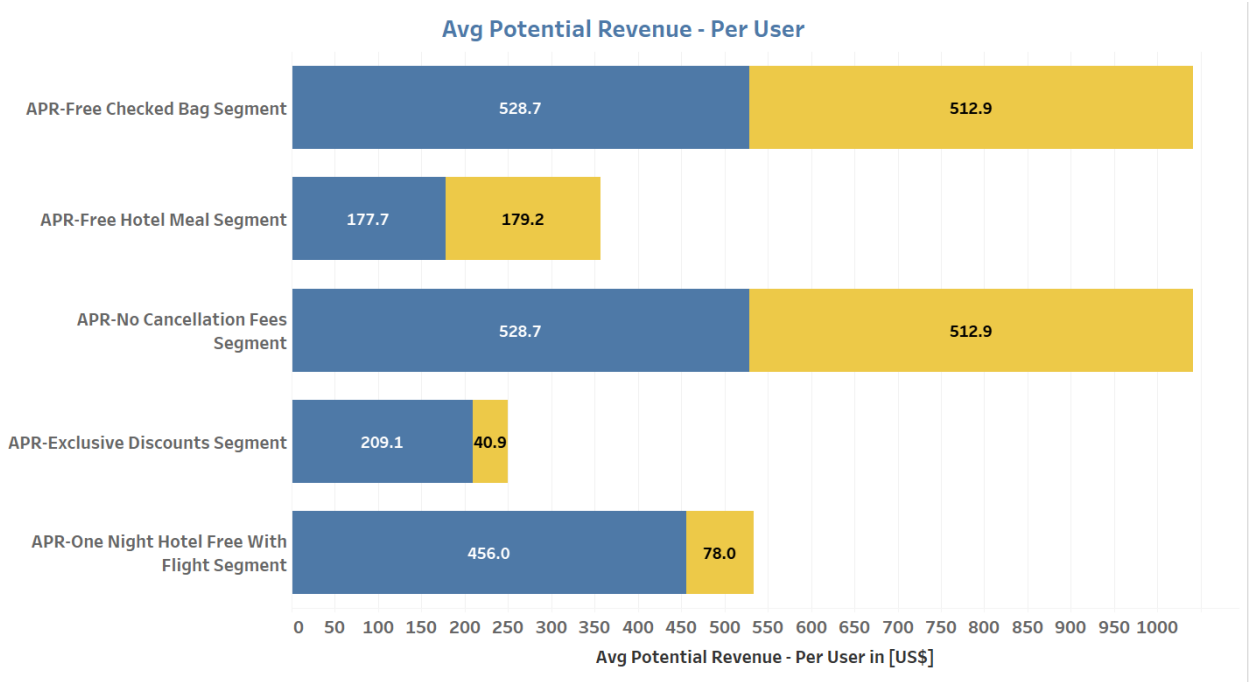
The **logic** for segmentation is based on the following:

- If eligible for more than one segment, assign to the first eligible lowest-cost segment
- If not eligible for any segment, assign to lowest-cost segment

Appendix 2-Customer Segmentation + Revenue Potential







Appendix 3-Reference links

- Video presentation: loom
- Tableau public Story: TravelTide customer segmentation + Revenue Potential