

Segmentation Analysis of TravelTide

By: The Analyst at TravelTide

Summary

This report aims to analyze the customers of TravelTide in their preferences of perks proposed by our Marketing team. In addition, we will assign each customer in the cohort of users as defined by Marketing with a likely favorite perk.

Context

TravelTide is a startup in the business of online travel booking. Its customers have access to the largest travel inventory in the e-booking space. To retain and add value to existing customers, the CEO Kevin Talanick brought in Elena Tarrant, the new Head of Marketing, to supercharge a new personalized rewards program. This strategy is based on understanding of customer behavior. And is proven to generate repeat business if executed well.

Elena has proposed five perks that are most likely favored by TravelTide's customers. As subject matter experts, our task is to apply data skills to validate the perks and assign customers to segments according to what we think their favorite perk is.

Results

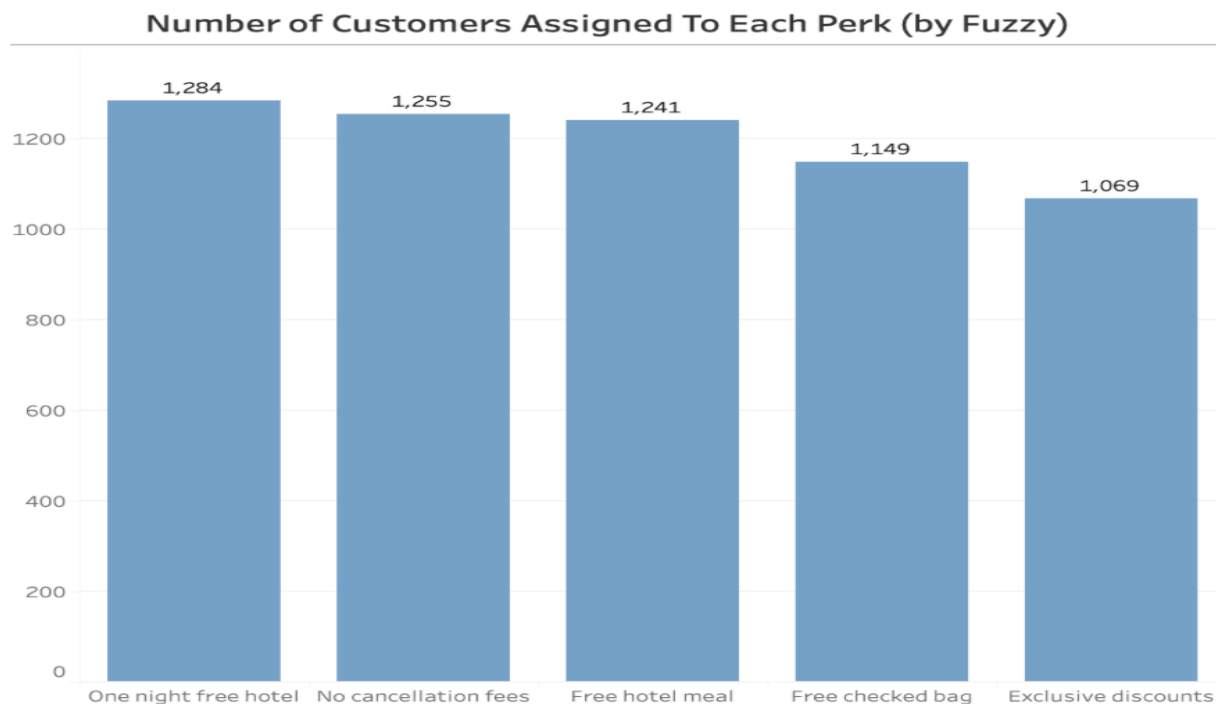
We use tools like SQL, Tableau, and Python to help with our analytic works performed against the TravelTide databases. We come up with several indexes that track customer behaviors in order to identify which perks are most attractive to which customers.

- 1) Free Hotel Meal - we combine metrics hotel discount proportion, average saving per night, and average hotel discount to create `free_hotel_meal_index`. The idea is to identify those customers who frequently book hotels with discounts. For those travelers, dine-in the same hotel they are staying offers convenience and comfort. They are likely to find this perk more appealing.
- 2) Free Checked Bag - the `avg_checked_bags_index` is based on the metric average checked bag per flight. It helps to identify those customers that booked flights with checked bags. Likely they would welcome a free checked bag as reward in their next bookings.

- 3) No Cancellation Fees - we calculated the rate of booked flights or hotels that were canceled. It helps us to look for those customers with cancellation history. They may find the free reward attractive when they book their next hotel or flight.
- 4) Exclusive Discount - we combine the metrics flight discount percent, average flight discount, and average dollar saving per km to create bargain_hunter_index. We are looking for those customers with flight discounts history. It indicates those customers actively looking for deals in their bookings. Exclusive deals should pique the interests of this group of customers.
- 5) 1 Night Free Hotel with Flight - we calculate the proportion of bookings that included both hotel and flight against all bookings to create the flight_hotel_booked_index. This helps us to identify those customers who often book hotels and flights together. This group of customers surely like the prospects of having free perks in their combination package.

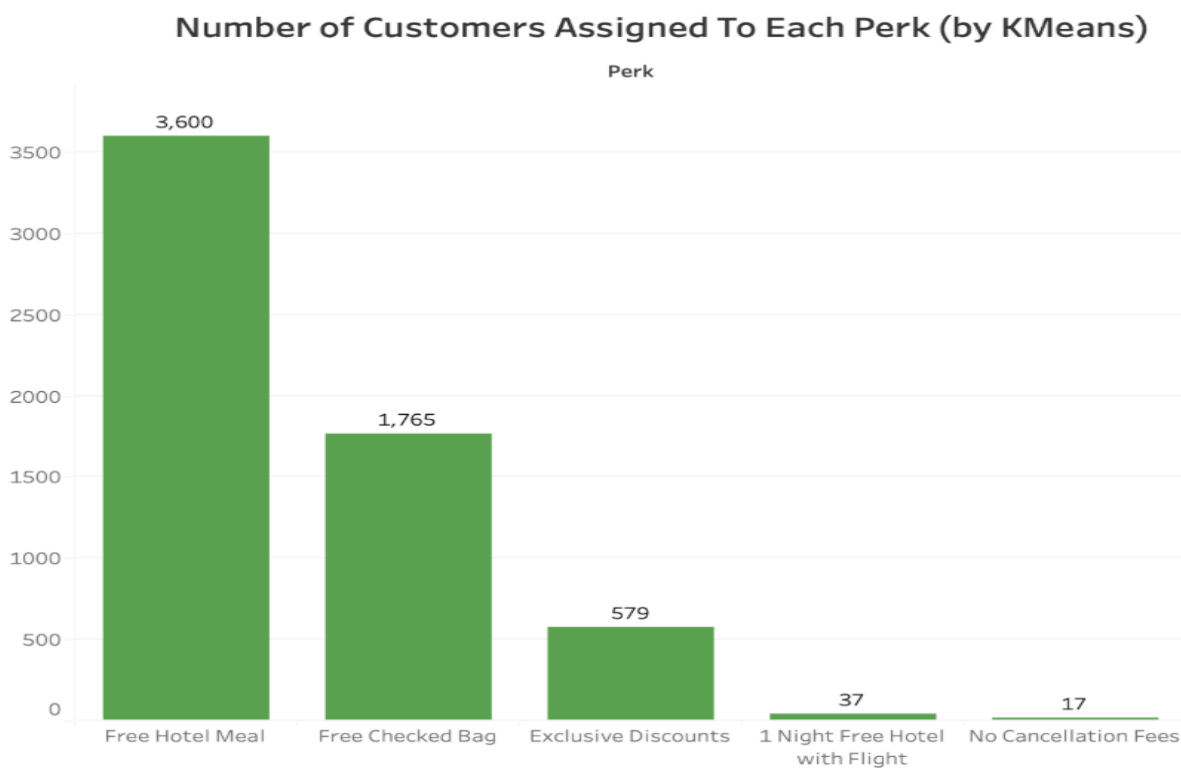
We applied two approaches to segment customers by perk types: Fuzzy Segmentation, K-Means algorithm.

In Fuzzy Segmentation, customers can belong to more than one segment after different indexes are computed. Then each customer is going through a ranking process where a perk they most belong to is assigned. From the following plot we can see how the perks are distributed across the cohort of customers.



The number of customers assigned for each segment does not vary much. For instance, the number of customers for the perk 'One night free hotel' (highest count) is 1284, versus the next highest count of the perk 'No cancellation fee' is 1255. Even the lowest count of customers of the perk 'Exclusive Discounts' has 1069, or 115 customers less than the perk 'One night free hotel'.

In the K-Means approach, customers are assigned based on the distance calculation between the center of a segment and its associated data points. The algorithm then assigned a customer to a segment at the end of the process. The following chart shows the perks distribution of the K-Means method.



The above shows the huge differences between the number of customers across segments. For example, the perk 'Free Hotel Meal' has the largest number of customers of 3600, and the second highest perk 'Free Checked Bag' has 1765 customers. The perk 'No Cancellation Fees' has the lowest count with only 17 customers.

Recommendation

The perks distribution charts from the previous section show the results generated by the two approaches we described here are not aligned. The difference can be attributed to the design of the algorithms.

In the Fuzzy method, we define the metrics and indexes in a top-down manner, meaning we apply the knowledge we have about the data and their affinity with the customers' behaviors. We can interpret the results in language both experts and non-technical audiences can relate to.

The K-Means method is based on rigorous statistical theory that can be used in many applications. It processes data in a bottom-up fashion, mainly through exhaustive distance calculation between cluster centers and associated data points. We need to find the right parameters (for example, number of clusters) and conduct comprehensive analyses to interpret and uncover meanings of final results.

The Fuzzy method is simpler to analyze and easier to interpret the generated results. Thus, we recommend using this approach to assign perks to our customers in the rewards program.

References

Link to Dashboard [TravelTide Story](#)