Nicholas Jones: TravelTide Segment Analysis (Final Revision)
09/18/2023

Introduction/Background

The main idea behind this segmentation analysis was to find an effective way to increase customer retention. An effective framework for this analysis was a perk rewards program; in particular, we chose the following 5 perks that all customers might fall into: Free Hotel Meal, Free Checked Bag, No Cancellation Fees, Exclusive Discounts, and 1 Free Night at Hotel with Flight.

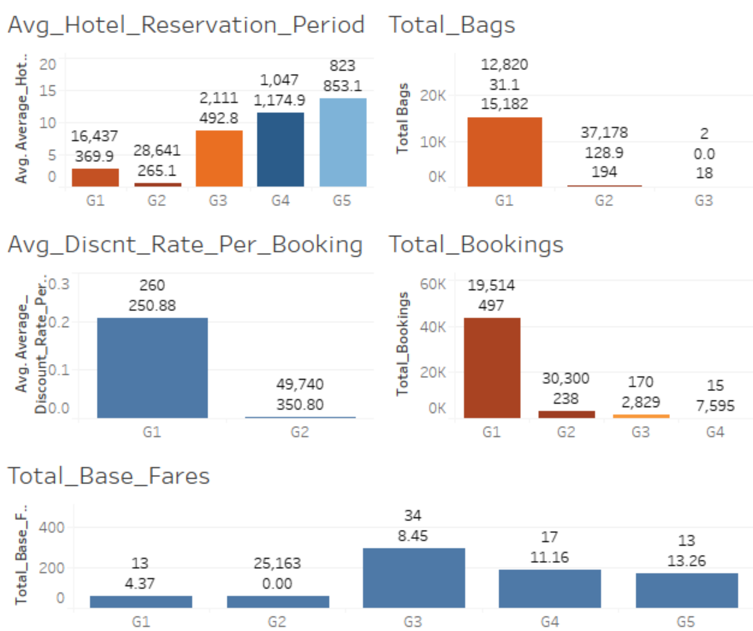
Objective/Summary

Based on a certain number of customer segments, the main objective is to find the characteristics of the customers that would find the perks very relevant and helpful; so much so, customers will become continuous customers and sign up for the rewards program.

Methodology

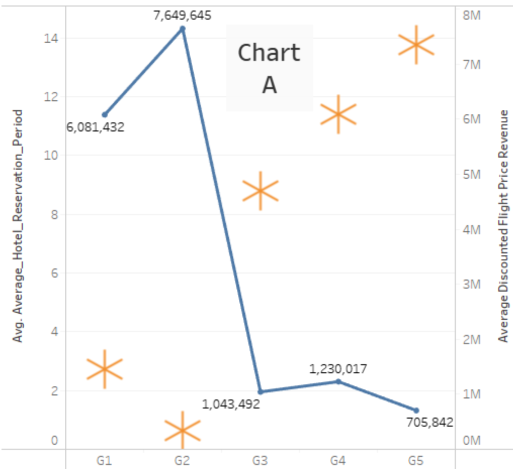
The following metrics are the ones that I chose for this analysis: **Average_Hotel_Reservation_Period** (average amount of days that buyers occupied a hotel room), **Total_Bags** (total sum of checked bags), **Total_Bookings** (the sum of the booked seats on a flight and booked hotel rooms), **Average_Discount_Rate_Per_Booking** (the midpoint of the 2 discount rates divided by the total number of hotel and flight bookings), and **Total_Base_Fares** (sum of the base fare prices). As far as the segmentation, I chose to group customers by using the mean and the margins of error of the average number of vacation days (**Vacation_Duration**): I created a segmentation by taking the interval that started with the average duration of a vacation (average = 2.63404 days) and ended at the point which was 3 margins of error greater than the average (average + (3)*4.13406 days); after this, I subdivided this interval into 5 equal subintervals with these conditions: "G1" (the group that is closest to the average), "G2" (the group that is the second closest group to the average), and likewise for "G3", "G4", "G5"- there is an "Outliers" group as well. Lastly, I tried to confirm my findings with an alternative model as well. The SQL code and Python code can be found in Appendix #1, and there is a "Top Perk" Dashboard that can visualize different segmentations of the data in Appendix #2.

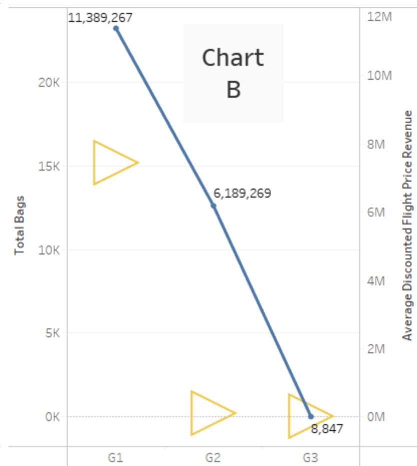
Key Findings (Part 1): Visualizations



<u>**Key Findings (Part 2): Regular Model**</u>

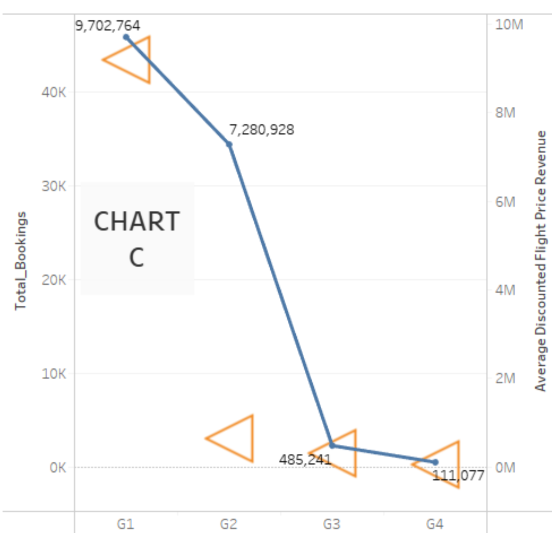
- For the Free Hotel Meal perk, I decided to look at the **Average_Hotel_Reservation_Period** metric: the customers that had vacation durations that are closer to the average should be the target. Also there are 45,078 customers in the "G1" and "G2", and all customers are in 5 groups. As a final note, the first chart below shows the metric values for each group and the approximate total flight booking revenue in USD (average discounted booking price for that group*the number of customers in that group) for each group. So, Chart A reveals that "G1" and "G2" yielded the highest revenues from flight bookings (the flight booking revenues are in the other charts as well).

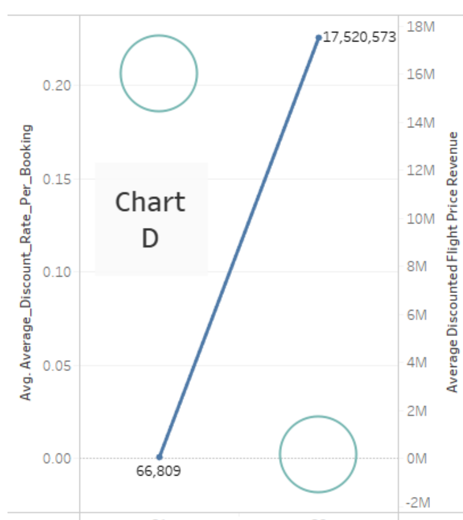




- For the Free Checked Bags perk, I decided to look at the **Total_Bags** metric: Customers with vacation durations that are closer to the average should be the target. Therefore, most of the customers remain in "G1" and "G2"(namely, 49,998), and all customers are in 3 groups. Lastly, even though the revenues are decreasing, "G1" and "G2" yielded the highest revenues from flight bookings, which can be seen in the line chart in Chart B.

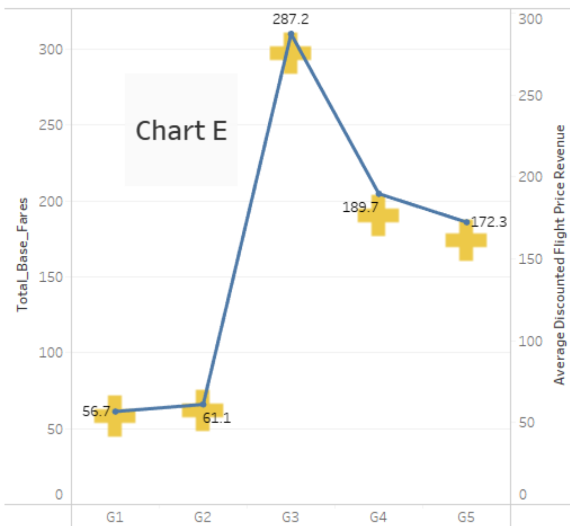
- To analyze the No Cancellation Fees perk, I decided to look at the **Total_Bookings** metric: The customers that had vacation durations that are closer to the average should be considered for the associated perk. Also, there are 49,814 customers in "G1" and "G2", and all customers are in 4 groups. Lastly, even though the revenues are decreasing, "G1" and "G2" yielded the highest revenues from flight bookings, which can be seen in the line chart in Chart C.





- To analyze the Exclusive Discounts perk, I decided to look at the **Average_Discount_Rate_Per_Booking** metric: The customers with vacation durations that are closer to the average should be the target group for this metric's associated perk. Furthermore, there are 50,000 customers in "G1"and "G2", and all customers are in 2 groups. Lastly, even though the revenues are increasing, "G1" and "G2" yielded the highest revenues from flight bookings, which can be seen in the line chart in Chart D.

- To analyze the 1 Free Night with Flight perk, I decided to look at the **Total_Base_Fares** metric: Customers with vacation durations that are closer to the average should be considered. Also, there is a sizable number of customers (namely, 25,210) that are in the "G1,""G2," and "G3" groups, and every customer is in one of 5 groups. Lastly, even though the revenues are increasing, "G1,""G2," and "G3" yielded the highest revenues from flight bookings, which can be seen in the line chart in Chart E.



Key Findings (Part 3): Alternative Model

- Although the alternative model did not produce a substantial amount of new information, it did yield one important new insight: The number of perks should be reduced to 3 because all of the customer data could be classified with that exact number of groups.

Recommendation/Suggestions

- The 2 models suggest that only 2 or 3 groups should be used. Therefore, I have concluded that Elena's model is not accurate. TravelTide needs to find 3 smaller groups, and this can be done by making that adjustment to my strategy.

Appendix

(1). SQL code file (file name: sqlcode.txt)/Python script (file name: MS_segment_final.ipynb):
[https://github.com/Ghubby30/MasterSchoolSegmentation]
(2). "Top Perk" Dashboard:
[https://public.tableau.com/views/Top_Perk_Dashboard-Revision/TopPerk?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link]