

# Customer Segmentation and Rewards Program Analysis

## Introduction

This analysis aimed to segment customers into meaningful clusters to tailor a rewards program with specific perks. The original perks considered were:

- Free hotel meal
- Free checked bag
- No cancellation fees
- Exclusive discounts
- 1 night free hotel with flights

## Methodology

- Using SQL to extract raw data from key tables and aggregate data.
- Processed and analyzed in Python to create 5 unique indexes as per each original perk.
- Used fuzzy segmentation to classify each user based on their affinity for each perk.
- Used K-Means clustering to check the engineered features.

## Data Preprocessing

- Handling Missing Values: Missing values in certain features were filled with zeros.
- Outliers: Outliers were handled by capping them to the 1st and 99th percentiles, retaining most of the data.
- Feature Scaling: Standardization was applied to make the features comparable.

## Feature Engineering

- Created new features representing eligibility for hotel meal offers, average checked bags per flight, cancellation rate, average discounts per booking, and a rate for booking both flights and hotels together.

## Exploratory Analysis

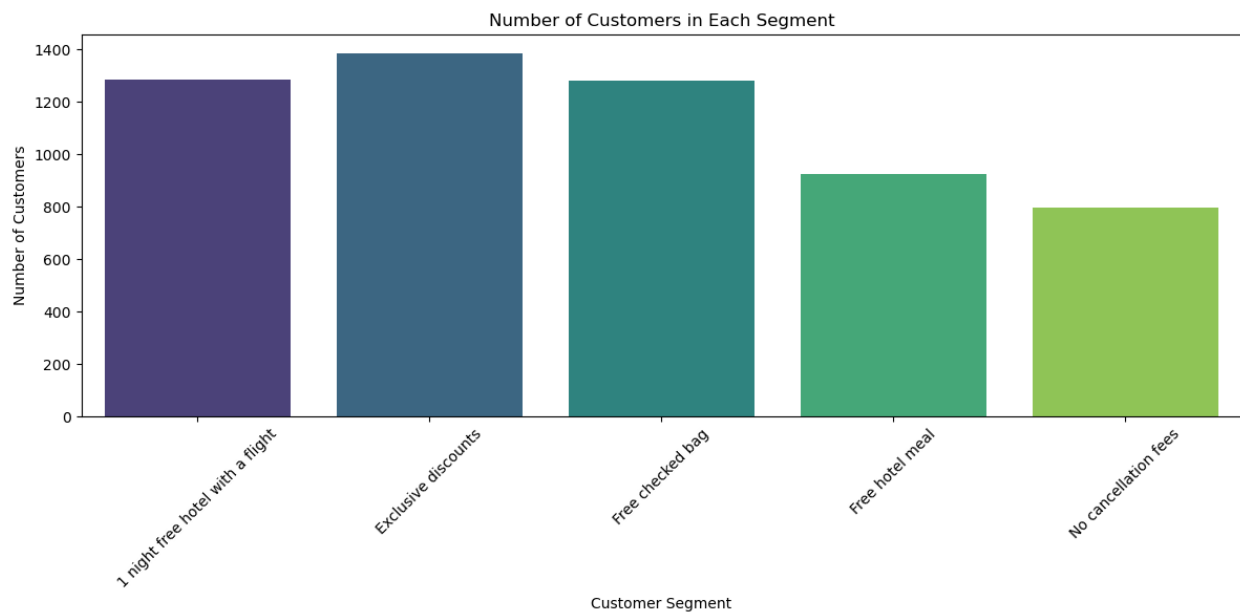
- Visualized distributions, relationships, and patterns within the data.
- Identified key characteristics and behaviors of customers.

## Clustering

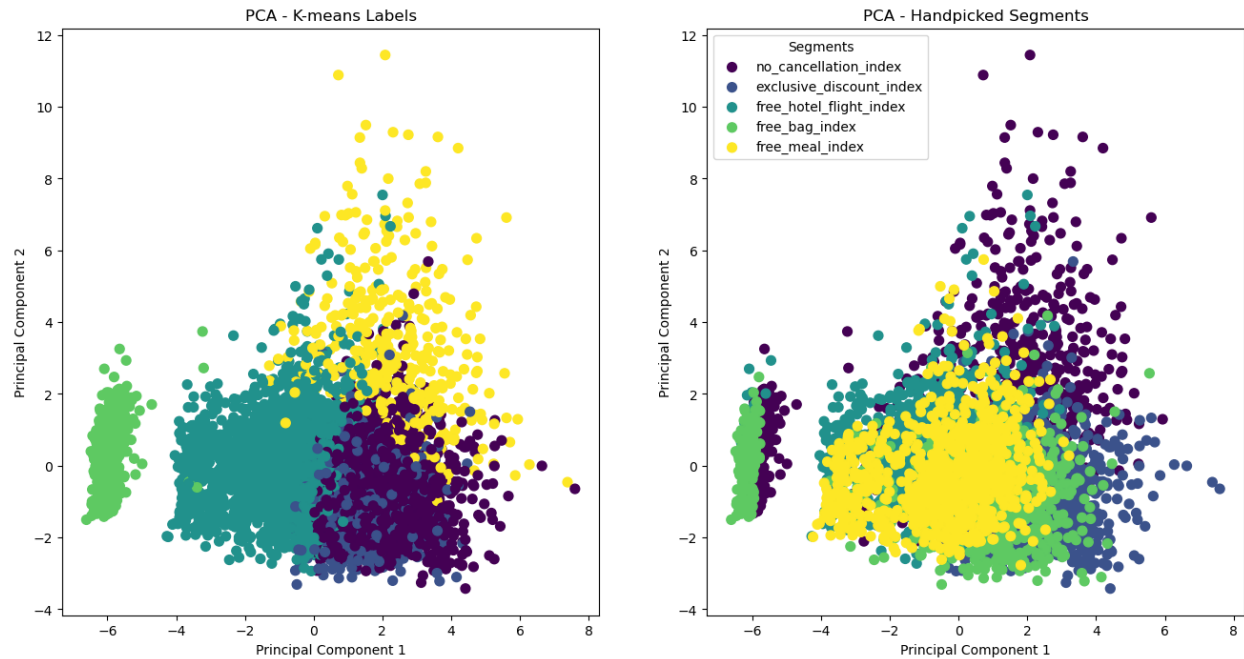
- Applied K-Means clustering with 5 clusters.
- Named clusters based on their characteristics:
  - Frequent Travelers: Often book both flights and hotels.
  - Discount Seekers: Seek flight and hotel discounts.
  - Occasional Travelers: Travel occasionally.
  - Infrequent Users: Rarely use services.
  - High-Spenders: High spending and cancellations.

## Key Findings

The segment with the "1-night free with a flight" appears to be the most popular, with the number of users ~1300. The "No cancellation fees" segment is the least popular, with the number of users ~800. Overall, there is a significant user base that would appreciate these reward perks.

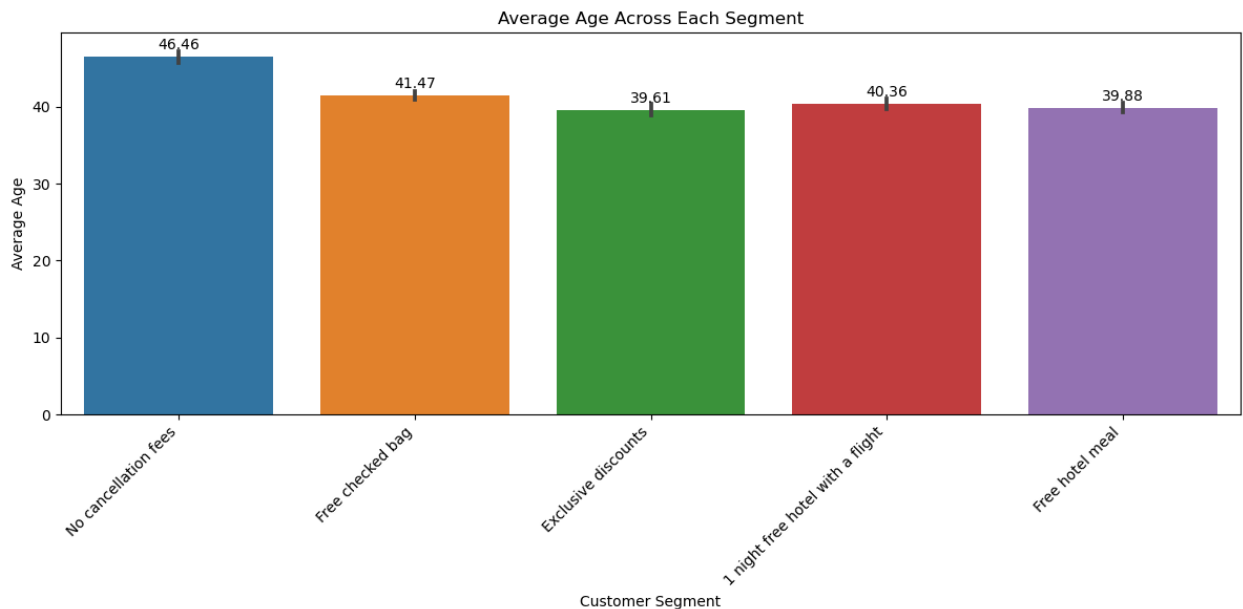


Each user was assigned to their highest affinity for each perk. Based on the K-Means clustering and engineered features, there is a clear enough separation for interest in each perk, though multiple users may be interested in other perks as well. This is an excellent opportunity for cross selling different perks.



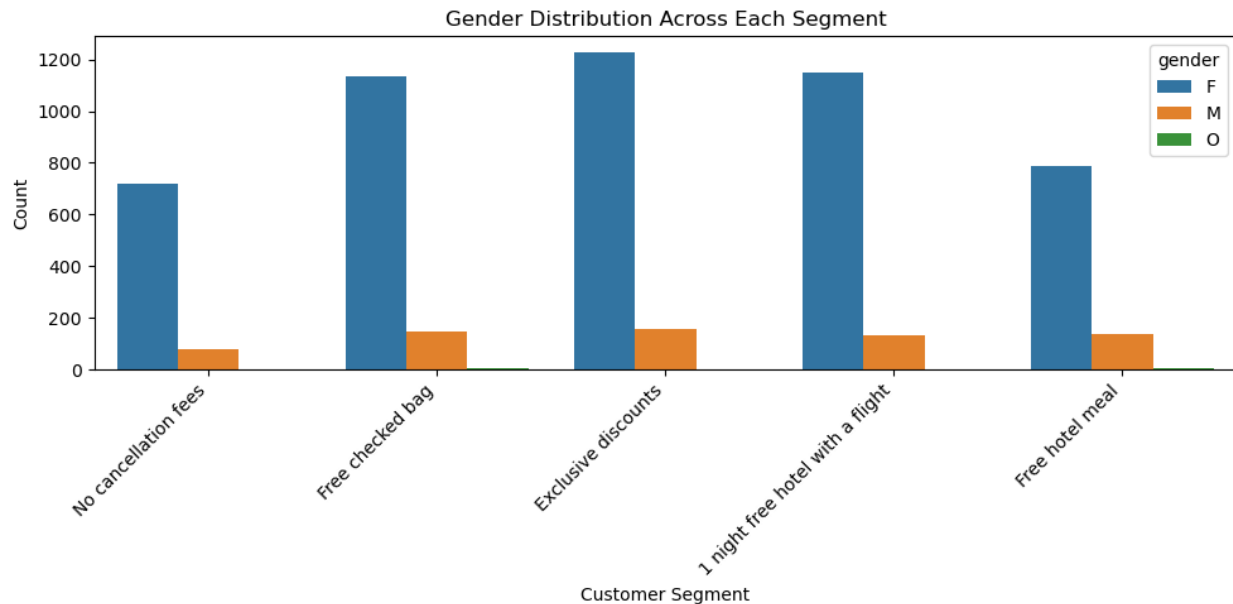
## Age

The chart suggests that older customers tend to value the option of no cancellation fees more compared to other benefits, which may indicate a preference for flexibility or risk aversion. Younger customers appear to be more attracted to exclusive discounts, which could imply sensitivity to price or a greater interest in getting the best deals. The relatively close average ages for the other benefits suggest that these offers appeal to a broad age range without significant variation.



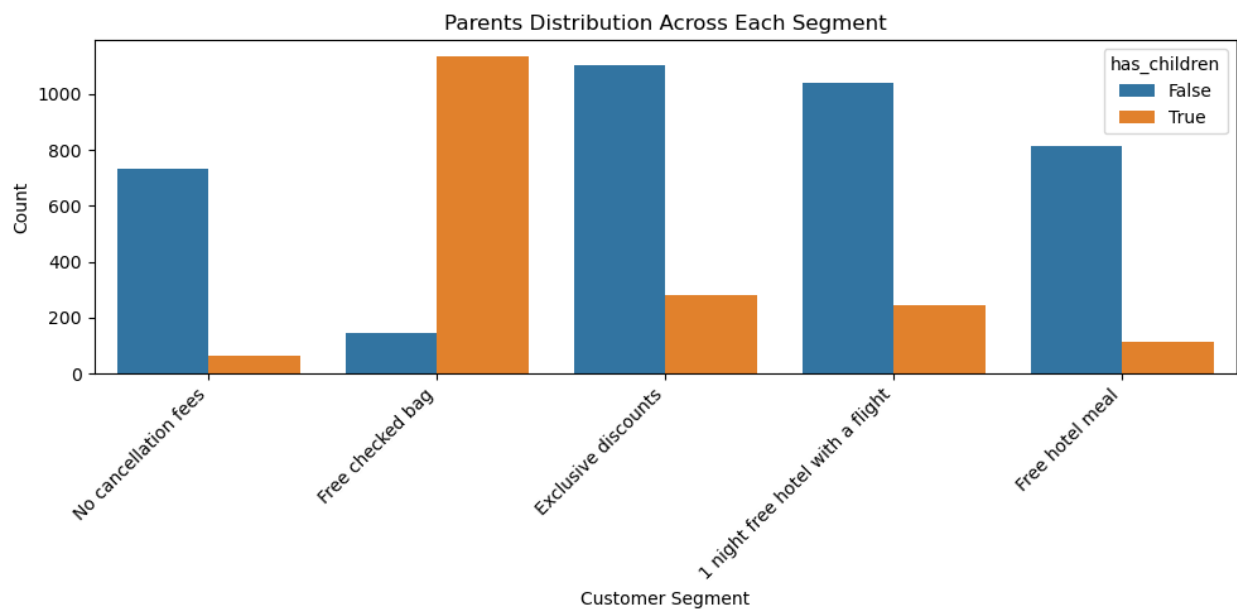
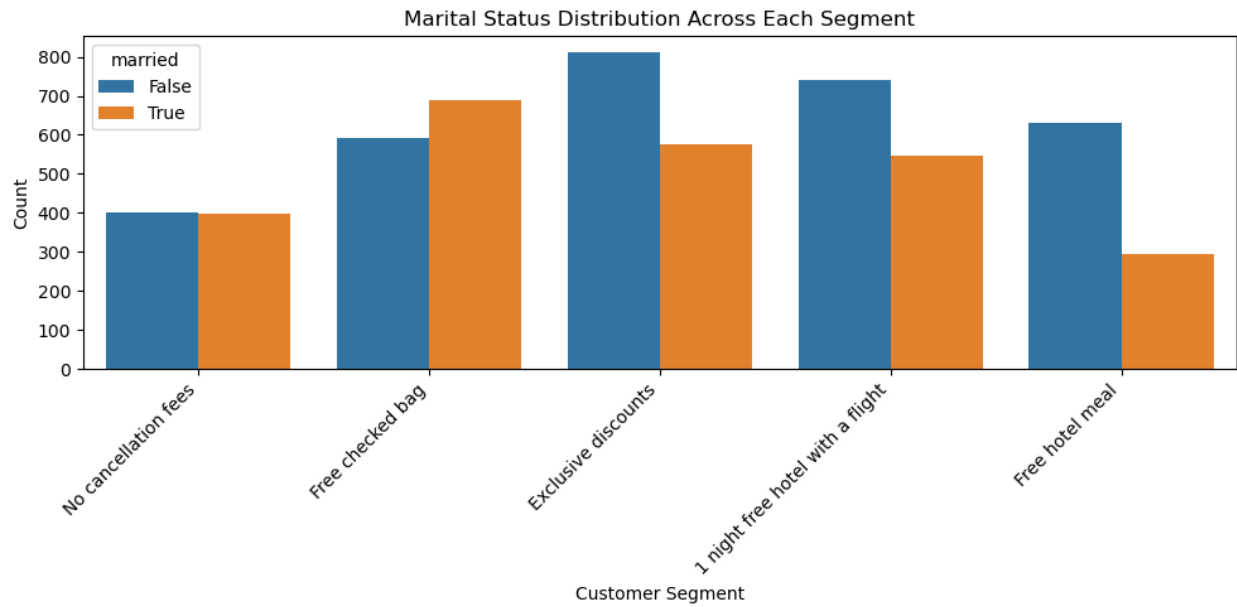
## Gender

There is a significantly larger female population in all segments. The features and marketing are particularly appealing to women. The number of males does not vary significantly across all segments, though the lowest representation is in No cancellation fees. The Exclusive discount segment has the highest female representation, suggesting this feature is most appealing overall to women.



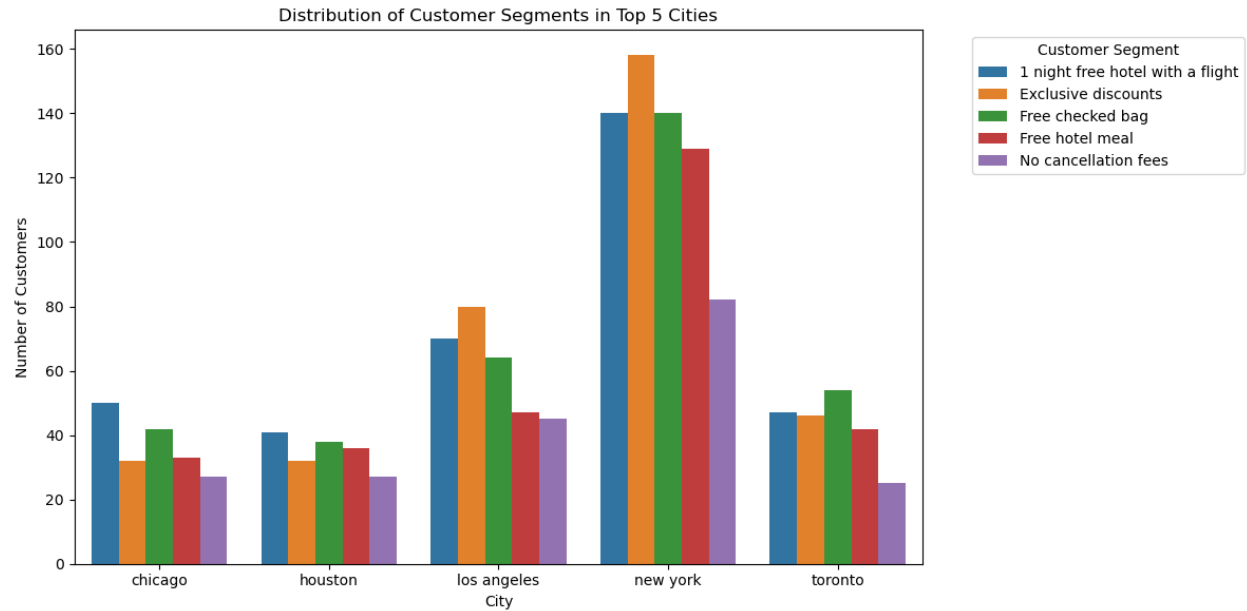
## Families

Interestingly, users who are married enjoy the benefit of free checked bags, whereas non-married users gravitate more to Exclusive discounts. Overall the rewards perks tend to be more interesting for users without kids, though the users who do have kids will likely take advantage of Free checked bags.



## Geography

Across all of the top 5 cities, we can see a keen interest in Exclusive discounts, especially in New York and La. Chicago would take advantage of 1 night free hotel with flight along with Houston. Toronto has a propensity to lean toward free checked bag.



## Disclaimers

- **Outliers Handling:** Outliers were capped, which might affect the representativeness of extreme cases.
- **Cluster Stability:** The clustering method and number of clusters were chosen based on analysis, and variations might lead to different results.
- **Ethical Considerations:** All analysis complies with relevant legal and ethical guidelines.

## Conclusion

The analysis successfully segmented customers and provided insights to tailor a rewards program with specific perks. The segmentation aligns with customer behaviors and preferences, offering a robust foundation for targeted strategies.

## Recommendations

- The Rewards Program can be launched with Exclusive discounts, 1 night free hotel with flight, and free checked bag: the largest segments.

- A/B testing the perks will help in understanding effectiveness before activating the program.
- Rotate the perks every three-six months and track effectiveness, considering the overlap of interest for perks.
- Track and observe engagement, conversions, and lifetime value to measure the success of the campaign.

Appendix:

SQL queries [here](#)

Python analysis [here](#)