Customer Segmentation for Marriott Hotel Bookings

Nexus 5: Boram Gaudet, Hanyang Gu, Jiahua Jia, Srita Kothuri

Research Question

- How can Marriott’s online booking customers be segmented based on factors like average amount spent, total guests, and region?

- Why does this matter?: By understanding their customer segments, Marriott can tailor digital experiences, personalize offers, and optimize revenue strategies for each group of customers

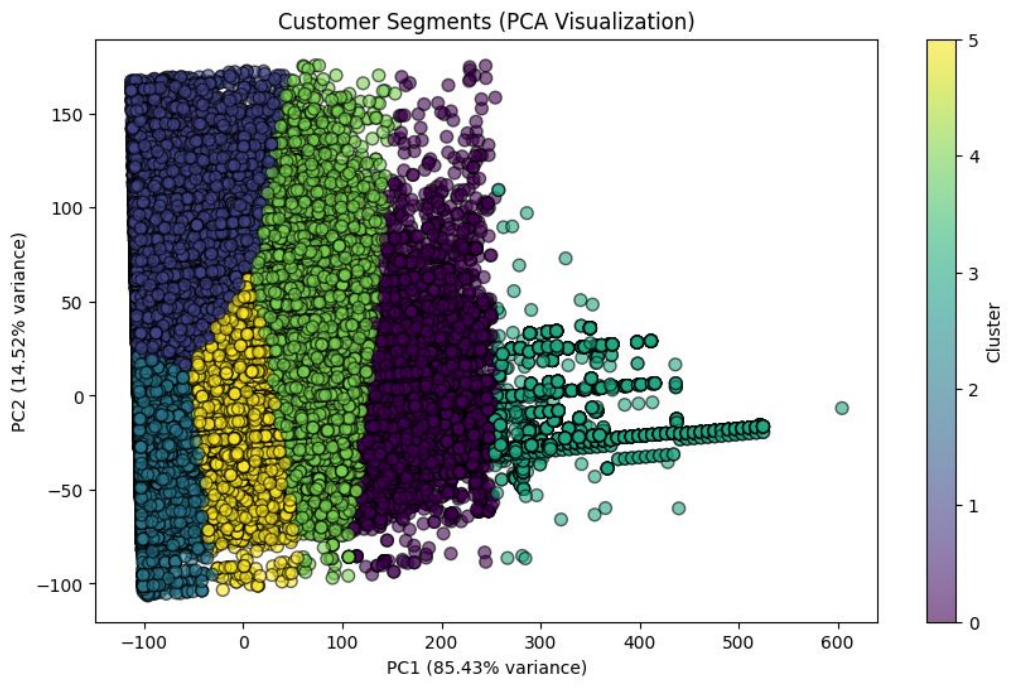
Data Description

- We are using the *Marriott Mobile & Online Booking Dataset* from Kaggle. This is a synthetic dataset of 119,390 Marriott hotel bookings

- There are 32 columns of variables, but we used only a subset of variables for our analysis:
  - Lead Time
  - Length of Stay
  - Total Guests (sum of adult and children guests)
  - Daily Rate
  - Season of booking
  - Region of booking (Asia, Europe, Americas, Oceania, Africa)

- After cleaning the data we were left with 116,347 unique values

Analysis



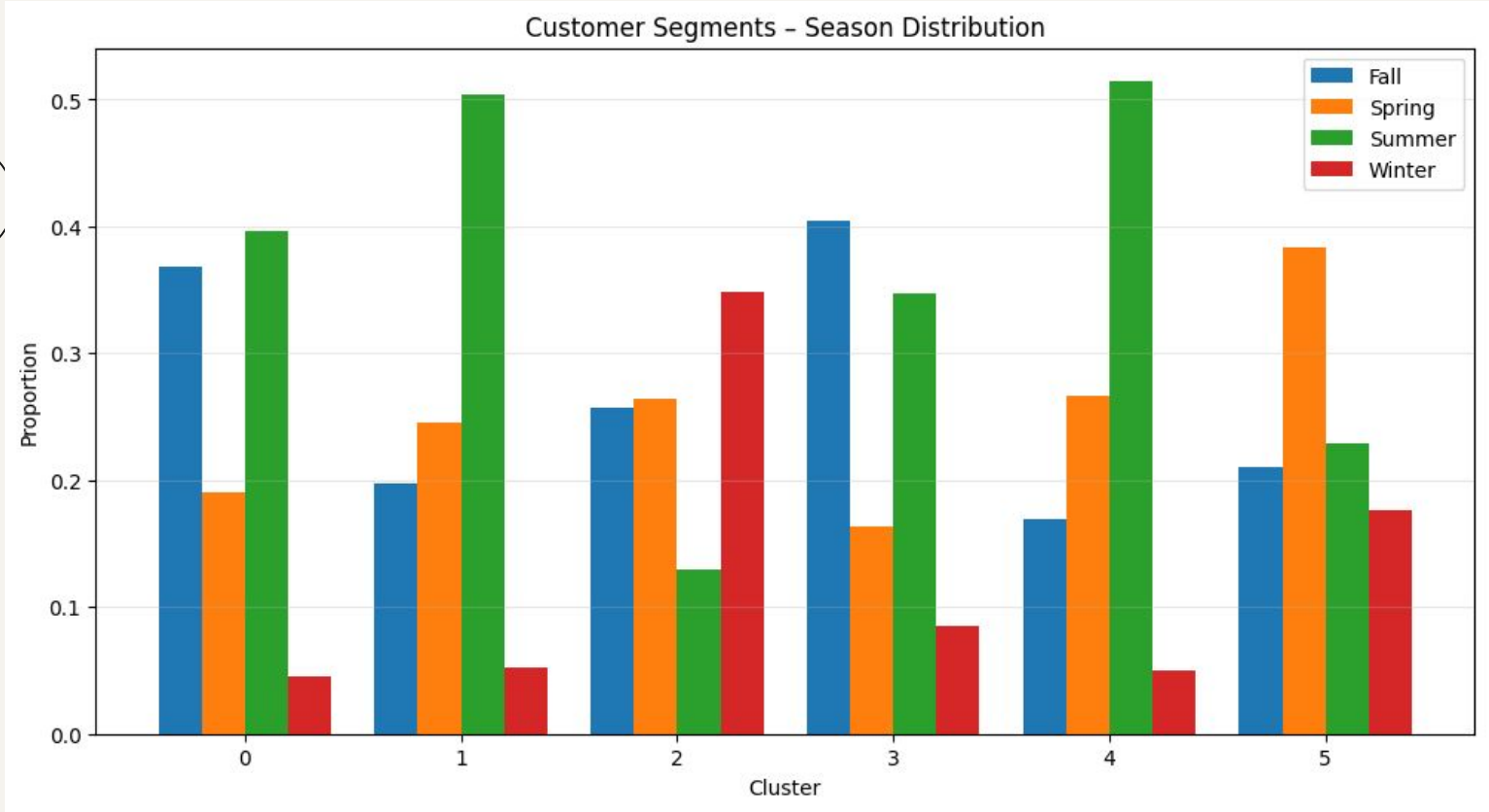
- Utilized Python to create K-Means clustering of customers based on key variables
- Apply standard scaler to numerical variables
- Set dummies for categorical variables
- Used Principal Component Analysis (PCA) to reduce the number of dimensions
- Identified and visualized 6 clusters of customers

Note: Optimal number of clusters is determined by Elbow Method

Insights

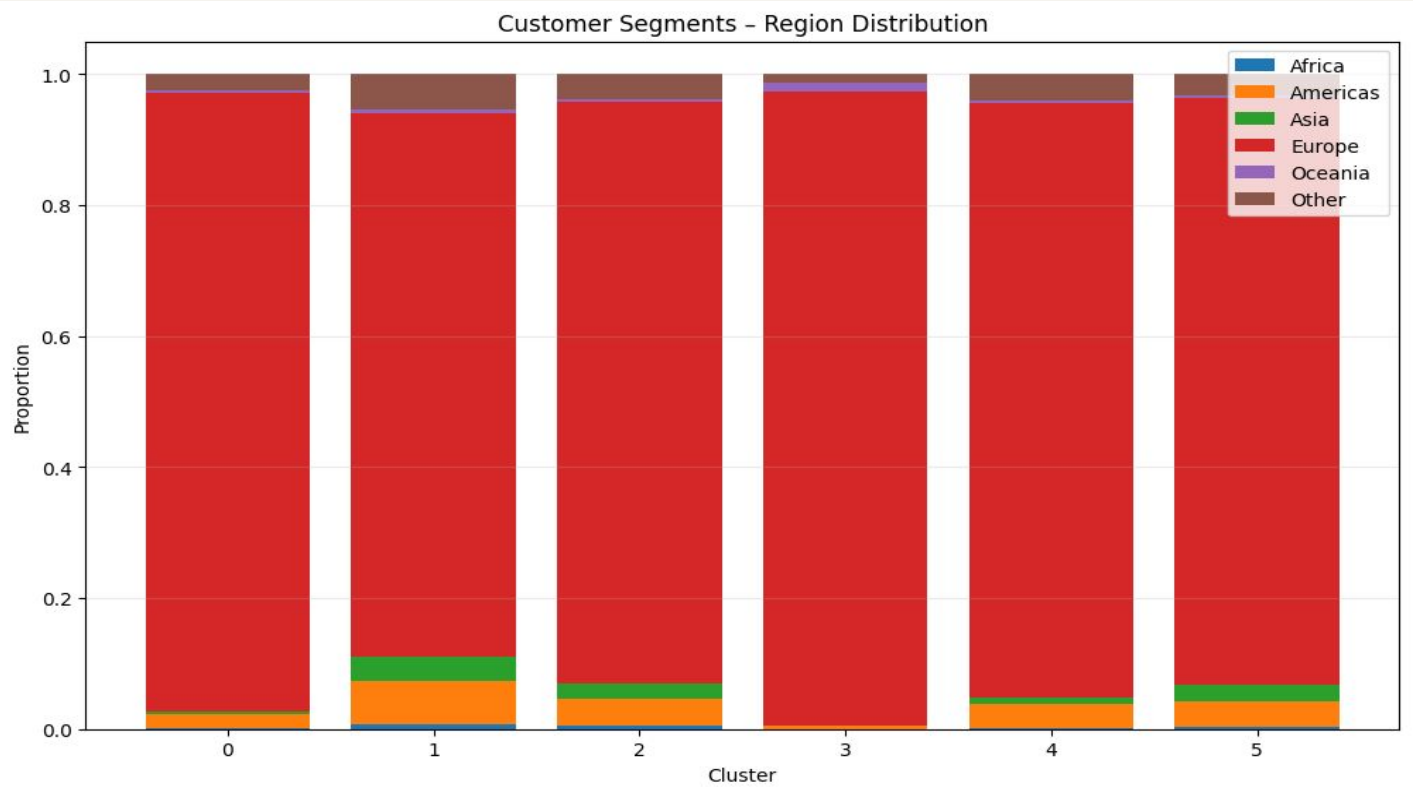
- To understand how each cluster differs, we compared them along all the key attributes
  - Total Guests
  - Season of booking
  - Region of booking
  - Lead Time
  - Daily Rate
  - Length of Stay

- All numerical variables were averaged to get a representative number for each cluster



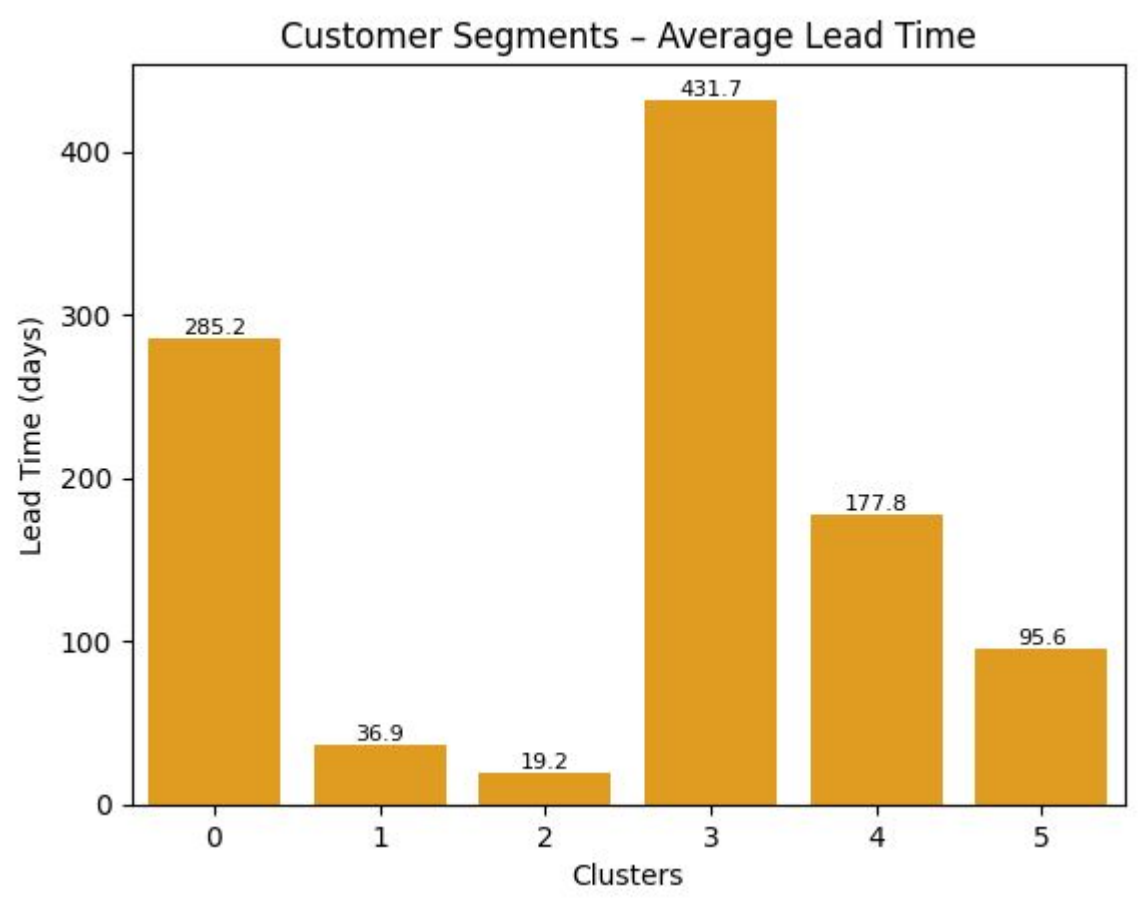
Key Findings:

- Clusters 0, 1, and 4 booked rooms mostly in the summer



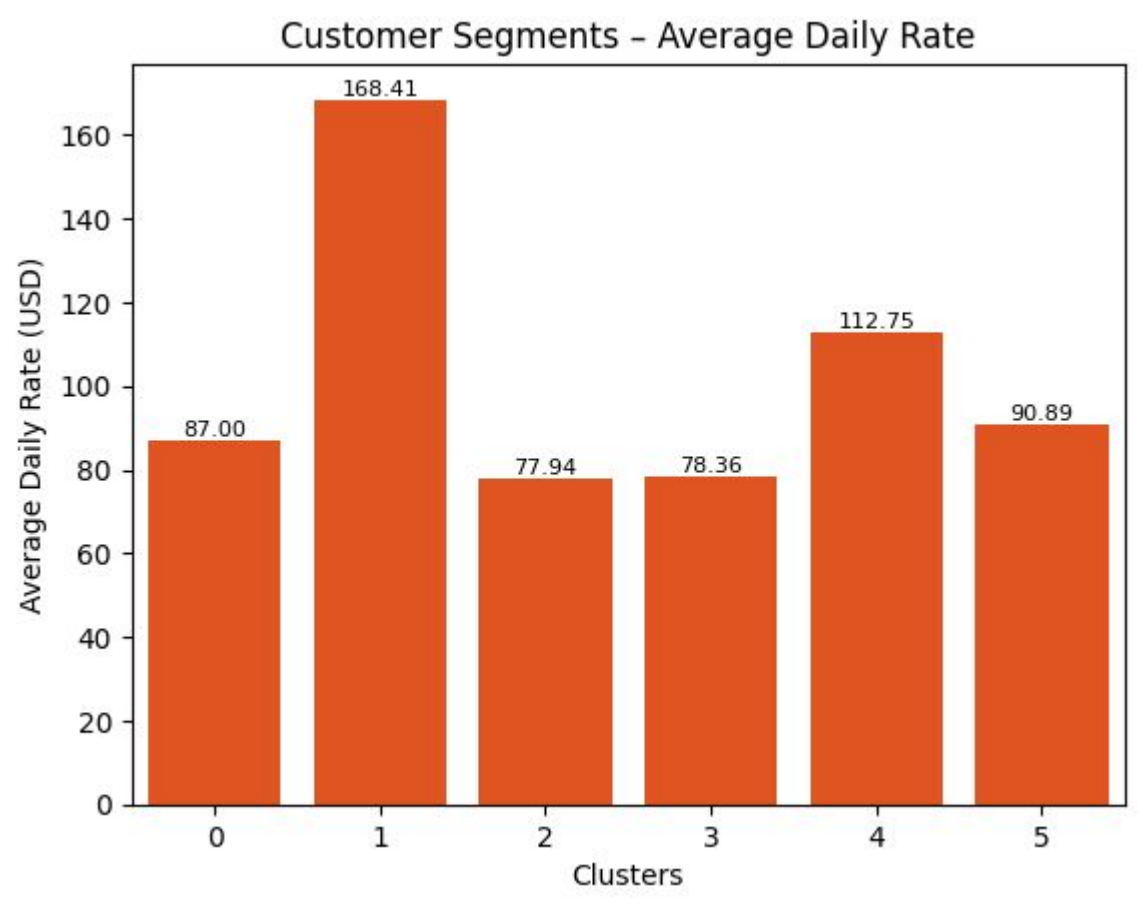
Key Findings:

- Most all bookings in the data set are in Europe (~10,3740; 89.16%)
- Segment 1 is the most varied



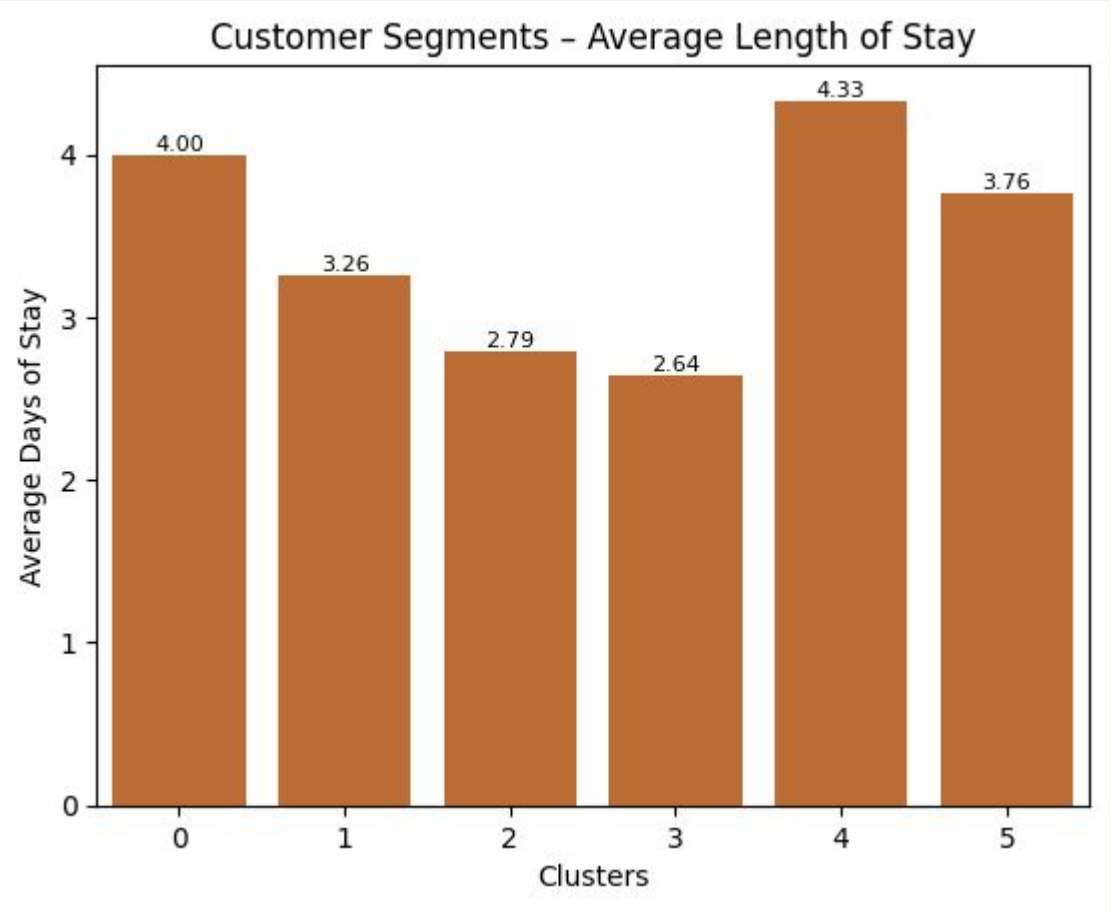
Key Findings:

- Cluster 3 had the longest lead time
- Cluster 2 had the shortest



Key Findings:

- Cluster 1 spent the most on their booking
- Cluster 2 spent the least on their booking



Key Findings:

- Cluster 4 booked the longest stay
- Cluster 3 booked the shortest stay

Summary of Customer Segments

- Cluster 0: Early Planners:
  - Plan far in advance and stay for a moderate amount of time
- Cluster 1: Business Travelers:
  - Book on short notice but spends the most per night
- Cluster 2: Spontaneous Budget Travelers:
  - Pay low rates and stay briefly
- Cluster 3: Early Budget Travelers:
  - Book incredibly early, stay for short periods of time, and dont spend alot
- Cluster 4: High Value Loyal Vacationers:
  - Pay above average prices and stay a long time
- Cluster 5: Standard Leisure Guests
  - Average on everything: planning time, stay length, and price

Recommendations for Marriott

Customer Segment	Type	Recommendation
0	Early Planners	Offer early-bird bundles (room + breakfast or resort credits) to lock in bookings earlier
1	Business Travelers	Offer business-oriented perks (mobile check-in, lounge access, late checkout).
2	Spontaneous Budget Travelers	Push mobile-only promotional rates to capture last-minute traffic before competitors do.
3	Early Budget Travelers	Introduce “Book Early & Save More” packages targeted at price-sensitive planners.
4	High Value Loyal Vacationers	Provide targeted loyalty upgrades (double points for long stays).
5	Standard Leisure Guests	Drive loyalty program sign-ups to convert casual users into repeat customers.

Managerial Implications

1) Segment Specific Pricing
    a) By understanding how each customer segment behaves, Marriott can tailor pricing strategies that aligns with the needs of each group

2) Effective and Targeted Promotions
    a) Clear customer segments allow Marriott to deliver promotions at the right time and through the most effective channels, which reduces wasted marketing resources

3) Strengthening Loyalty Relationships
    a) By identifying customer behavior patterns, Marriot can identify which specific groups have the most potential in becoming premium long-stay guests

Limitations of Our Work

- Data sourced from a public website
- Data is synthetic
  - Although it *models* real-world problems, it does not represent actual real-world observations
- We only focused on customer segmentation
  - Future work could identify which variables are most important in predicting booking/cancellation for each customer
  - This will give us a more holistic understanding of our customers

Thank You!

Citation

Hotel Booking Data. Kaggle. https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand