

BANA 200: Group Project

Group 8B : Made Jessica Felicia, Fabián Cruz-Ruiz, Yi-En Ivy Liu, Grace Kim, Chieh-Yi Chou

Introduction

The purpose of this project is to assist Sun Country Airlines in maximizing the use of its customer data to more effectively inform the marketing and product development teams so that the small airliner can better navigate the competitive airline industry. As a small airliner, the company needs to have a deep understanding of its customer segments, habits, and demographics to avoid poaching from larger airlines.

To first assist the airliner, we analyzed the raw data in the provided Excel spreadsheets to see the type of information we would be working on and how we could best illustrate the data's different variables. After initial review, we then reviewed the Customer Data Dictionary to infer which variables we believed to be the most important in answering the given questions such as the customer's birthdate, booked class, booking channel, base fare paid, membership status, booked product, flight route details, group size, seasonality habits, and days for pre-booked flights. We identified these 10 variables as the most flexible in comparing and visualizing our customer segments. Before being able to visualize or compare our variables, we noticed that the data provided by Sun Country Airlines was good regarding the information collected per customer, however, the use of two separate Excel spreadsheets made it difficult to cross-examine customer information. To address this issue, we decided to combine both spreadsheets into one data frame through Python in Google Collab. From there we conducted K-means clustering to group the data into different clusters and identified that dividing our data into five groups was the most efficient way to analyze the data.

The K-mean clustering made the data much more usable and easier to read as we were able to identify the optimal number of clusters to be five. In order to name each cluster, we then applied our course learning within Python to create various pie charts, bar graphs, histograms, and statistical graphics. These charts were based on the ten important variables we had previously identified and used to extract unique insights about each segment.

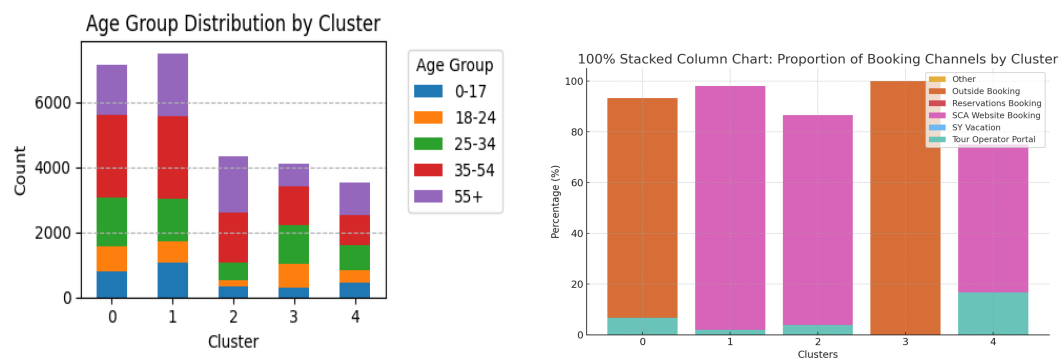
Through data visualization we were able to identify the key names of each segment as; Middle-Aged Standard Travelers (cluster 0), Middle-Aged Group Travelers (cluster 1), Senior Winter Escapers (cluster 2), Early Minneapolis Travelers (cluster 3), and Age-Blend Bargain Birds (cluster 4). After this, we created three unique visualizations for each cluster that are useful for Sun Country Airlines.

The Five Clusters & Their Underlying Key Variables

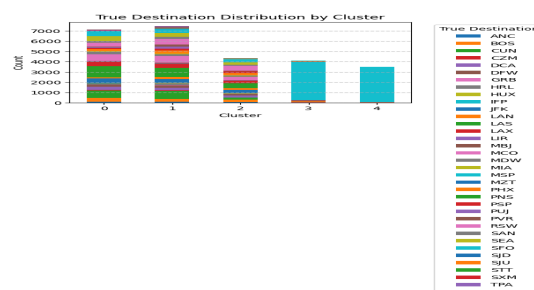
1. Cluster 0 - Middle-Aged Standard Travelers

Age Group: The largest age groups in Cluster 0 are **35-54** and **55+**, representing older, potentially more affluent travelers, suggesting that marketing efforts tailored to older travelers might resonate well with this segment.

Booking Channel: The vast majority of bookings in Cluster 0 are made through **outside booking** channels (over 7000 customers), with minimal use of **SCA website booking** or other channels. This suggests a strong preference for booking through third-party platforms.



Diverse Destinations: Customers in this cluster travel to a wide variety of destinations, with no single destination dominating the distribution. It also shows a pattern similar to Cluster 1.



(Refer to Appendix for detailed visualizations)

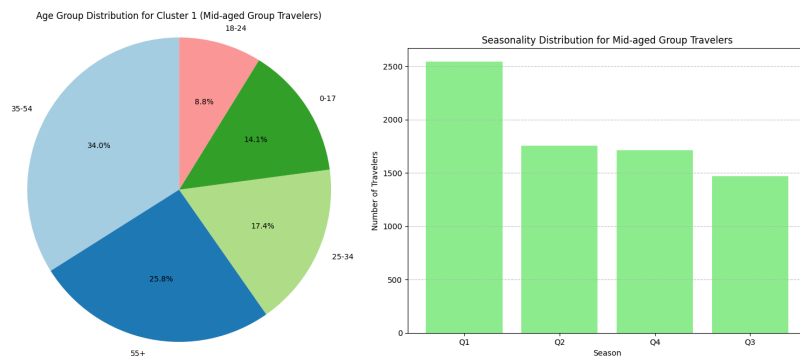
Marketing Insights:

- **Strengthen Partnerships with Third-Party Platforms** : Collaborate more closely with Online Travel Agencies (OTAs) and meta-search engines (e.g., Expedia, skyscanner, Google Flights) to promote Sun Country flights. Negotiating better visibility on these platforms, such as premium listings or advertisements, can increase exposure.
- **Review and Reputation Management** : Since many travelers rely on **reviews** from third-party booking platforms, it's important to maintain positive reviews and a high rating. Encourage satisfied customers to leave feedback and actively manage responses to both positive and negative reviews to enhance brand trust.
- **Destination-Based Custom Packages** : Both clusters show a preference for certain key destinations (e.g., BOS, DFW, LAX). Sun Country Airlines can create **destination-specific packages** that bundle airfare with hotel accommodations, car rentals, or local tours. Offering discounted **flight + hotel** or **flight + car rental** packages can incentivize these travelers.

2. Cluster 1 - Middle aged group travelers

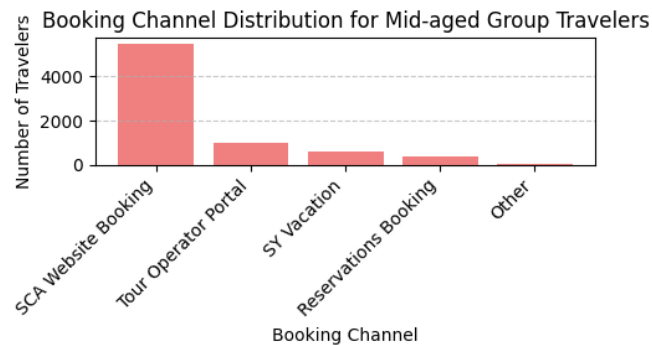
Majority are elderly: This group consists of more elderly travelers than other groups (both 35-55 & 55+), which suggests a need for services catering to seniors, such as assistance with boarding, comfortable seating options, and in-flight services that prioritize convenience and comfort.

Prefer to travel in Q1: This group prefers traveling in Q1 , likely taking advantage of post-holiday travel deals or spring break vacations. Targeting promotions during this period would be highly effective.



Have the largest average group size than other groups: They tend to travel in larger groups, potentially indicating family or group vacation plans. This is an opportunity to promote group discounts, family packages, or offers for tour groups.

- Mostly not members of the airline's loyalty program.
- Frequently book through SCA and On Tour Operator channels.

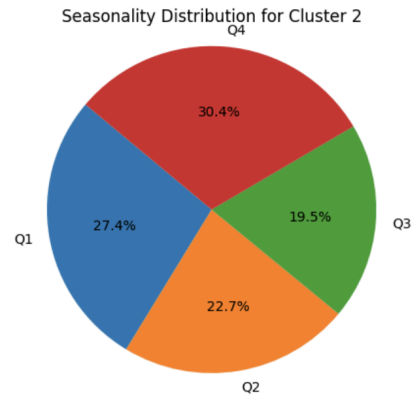
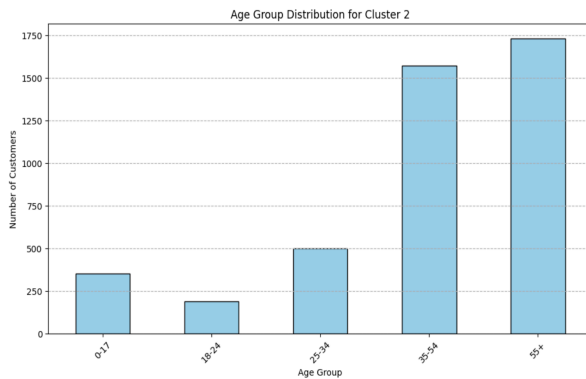


Marketing insights:

- **Promote Spring Travel Packages:** Offer discounted fares or special deals for winter travel that cater to larger groups. This could include family packages, multi-passenger discounts, or group travel incentives for seniors.
- **Bundle Offers:** Partner with tour operators to offer all-inclusive vacation packages, where flights, accommodations, and activities are bundled. Since this group prefers larger groups, these packages should be easy to book for multiple travelers.
- **Highlight senior-friendly services** such as priority boarding, wheelchair assistance, and enhanced customer service in marketing campaigns.
- **Targeted Campaigns for Ufly Membership:** Given that people in this cluster are not loyalty members, offer tailored incentives to join the Ufly program. This could include bonus points for group bookings, exclusive discounts for Q1 travel, or senior-specific perks.

3. Cluster 2 - Senior Winter Escapers

Cluster 2 has a key characteristic of people who have an age group of 55+ and most likely to travel in Q4 or Winter Quarter, this could indicate that the customer in this cluster is possibly traveling for Thanksgiving, Christmas holiday to visit their family.

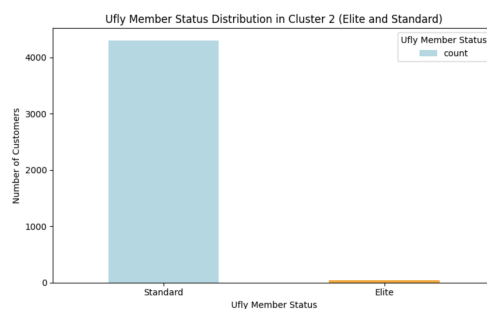


Average Group Size: With an average of 1.9, the customer in this cluster tends to travel solo or with a partner such as spouse or close friend.

Days pre-booked: They booked trips as advanced as 67 days in average, this could mean that they like to thoroughly plan and organize their trips.

Base Fare: As the mean base fare amount for this cluster is around \$302 which is the highest amount out of all clusters, indicating that the customers are willing to pay more for quality service, and comfort.

Ufly Member: This cluster consists of the most counts for standard Ufly Member, this could indicate that the customers are using their members to get benefits such as service preferences (in-flight service, seating policy, any other special accommodations) offered by Sun Country Airlines.



Marketing Insights:

- Highlight Comfort and Convenience as a Standard Ufly Member:** The customers in this cluster are willing to pay for comfort and convenience and are already engaged as a Standard Ufly Member. Sun Country Airlines could focus their marketing efforts on premium services, such as Ufly bonus points, extra legroom, priority boarding, special

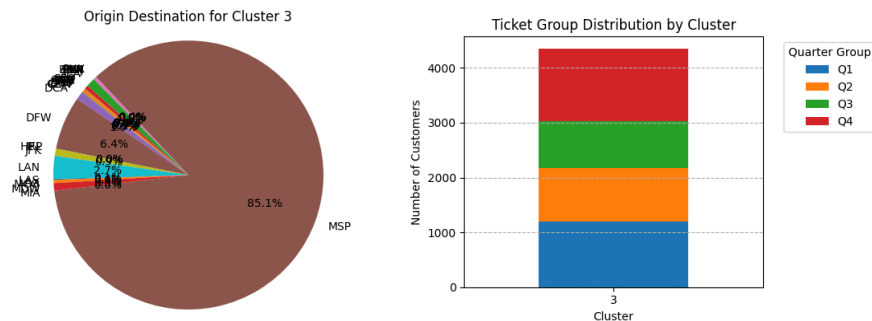
accommodation such as wheelchair assistance, or bundled packages that offer additional comfort to increase loyalty.

- **Promote Early Booking Discounts and Offers:** Sun Country Airlines can promote early-bird discounts such as “Fly Home for the Holidays” or “Reconnect with Loved Ones this Winter” targeting trips during the winter quarter (Q4). Sun Country Airlines could also leverage the option of flexible booking, meaning that the customers could have a free change flight that would appeal to their organized and cautious nature of traveling.

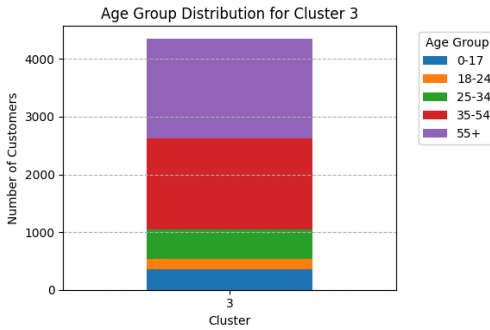
4. Cluster 3 - Early Minneapolis Travelers

Cluster 3 consists of 3838 travelers with over 85% of all them originating from one specific airport, Minneapolis-St.Paul International Airport (MSP). This is a key difference for this segment of travelers as the other 4 clusters are more diversified in terms of where travelers originate their travel from.

Seasonality Booking: This cluster’s customers are more likely to travel during the early months of the year, January through March, which is why we decided to name this cluster as “Early Minneapolis Travelers.”



Main Group: The main group of travelers within this cluster are coach flyers aged 35 to 54.



Base Fare: This cluster holds the second most expensive base fare when compared to the other four segments.

Marketing Insights:

- Since the average pre-booking date is 61 days before a flight, Sun Country Airlines should begin marketing efforts during early October and late December to target the “Early Minneapolis Travelers” early-year travel. This timing perfectly aligns with holiday shopping trends which can allow the airliner to offer promotional products tied with the near year. Some slogans the company can use are “Fly into the new year with this great offer!” or “Ring in the new year with a great deal in!”
- Since most of the travelers in cluster 3 originate from MSP, Sun Country Airlines could focus on offering amenities to these travelers such airport lounge access, complimentary Wi-Fi, or discount deals to retain them.

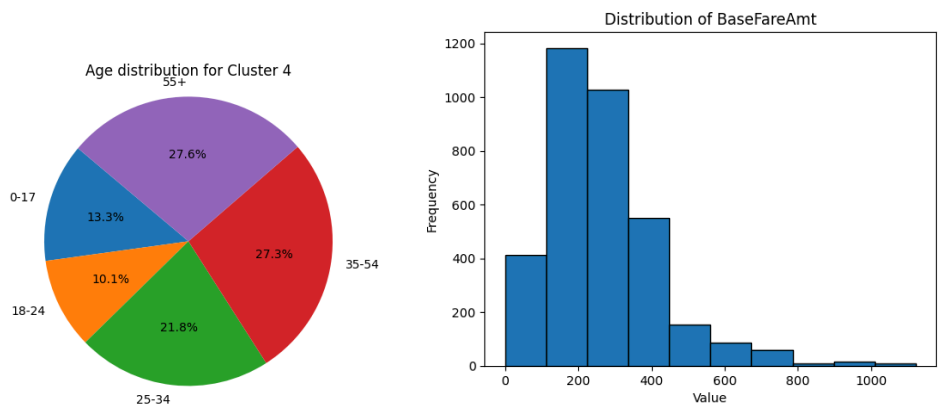
Risk Mitigation

- Not only does this cluster reveal potential market opportunities but also reveals an area where MSP airport-specific delay would impact the majority of travelers within the cluster. Considering this cluster is the second largest, it’s vital Sun Country Airlines maintains close monitoring of the airport’s operations and maintains plans for mitigating any potential airport disruptions.

5. Cluster 4: Age-Blend Bargain Birds

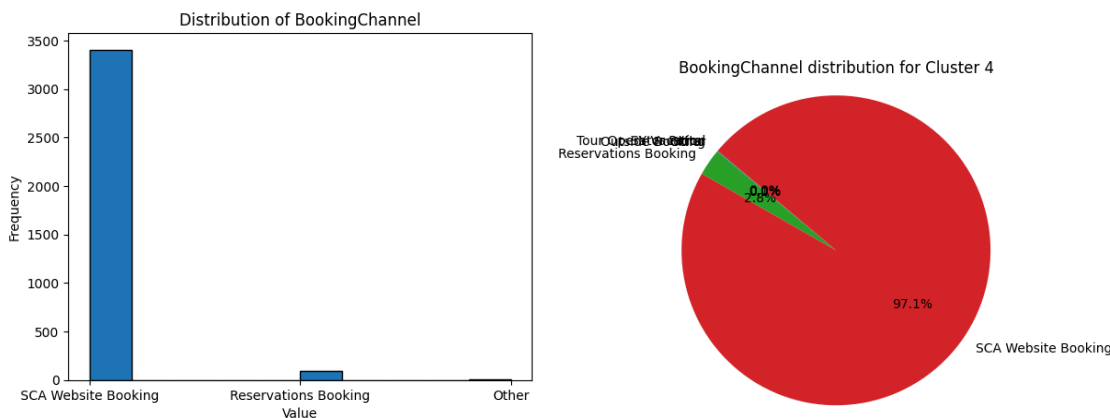
Cluster 4 has a key characteristic of people who have a balanced age distribution, demonstrating a diverse range of ages from young adults to seniors. This cluster also shows a strong preference

for booking flights through the SCA website and is particularly attentive to budget-friendly options, seeking the best value for their money.



Balanced Age Distribution: The age distribution within this group is relatively even, encompassing a wide range of age brackets. This balanced demographic suggests that the group is not skewed toward any particular age range, reflecting a broad appeal across different age groups.

Economical Fare Spending: The group exhibits a trend toward spending less on the base fare (BareFare) of their tickets. This indicates a cost-conscious approach to travel, where members are likely seeking value for money and might be inclined to choose budget-friendly options.



Booking Channel: Most customers in this group are using the SCA website when booking their tickets, indicating a strong preference for online reservations through this platform.

Marketing Insights:

- **Targeting:** To appeal to the Age-Blend Bargain Birds, marketing strategies could emphasize budget-friendly travel options, special promotions, and exclusive deals available through the SCA website. Highlighting the value for money and showcasing how travelers of all ages can benefit from affordable rates and convenient booking features will resonate with this diverse audience.
- **Communication:** Messaging should cater to a broad audience, ensuring inclusivity in terms of age and focusing on the benefits of early planning and cost savings. Use language that appeals to various age groups and emphasizes how planning ahead can lead to significant savings.
- **Potential Offers:**
 - ◆ Age-Based Discounts: Special promotions or discounts tailored to different age groups, such as senior savings or youth specials, to appeal to the diverse age range in this cluster.
 - ◆ Exclusive Online Deals: Limited-time offers available only through the SCA website, encouraging users to book directly through the platform.
 - ◆ Bundle Offers: Packages that combine flights with other services, like accommodation or car rentals, at a discounted rate, appealing to cost-conscious travelers looking for value.

Conclusion

By identifying the key traits of each customer segment, Sun Country Airlines can create more focused marketing campaigns. Each segment has unique booking habits, travel patterns, and group details, which lets the airline personalize promotions for different customer segments. Through data analysis, we found that using K-means clustering and visual tools can reveal important insights about customer behavior.

Clear and simple visualizations are essential for turning data into useful insights that support Sun Country Airlines' business goals. Each visualization should focus on key objectives, like finding growth opportunities or improving customer loyalty programs. Connecting these visual insights

to strategic actions, such as improving third-party booking partnerships, helps turn data into practical steps toward meeting business goals.

To effectively communicate these insights, we utilized a variety of visualization tools, ensuring our client gains a clear and intuitive understanding of their customer base. These tools include heatmaps, stacked bar charts, pie charts, histograms, and bar charts. Each of these visualizations was carefully selected to present data in the simplest, yet most informative way, avoiding any complexity that might obscure the message. For visualizations that introduced confusion or complexity (e.g. heatmap; refer to appendix below), we excluded them to maintain clarity. We focused on using the most effective tools to deliver insights tailored to each segment, enhancing the overall decision-making process for Sun Country Airlines.

Understanding and exploring the data from Sun Country Airlines' customer segments has proven invaluable in driving actionable insights. By analyzing key customer behaviors and traits such as booking patterns, group sizes, and fare preferences, we were able to uncover meaningful trends without relying on overly complex models. This approach allowed us to create clear, practical recommendations that align directly with the airline's business objectives. Over-complicating the process with advanced models could obscure the real-world insights that matter most. Instead, our focus on simple, straightforward data analysis ensures that the findings are both accurate and easy for executives like Warnken and Vaughan to implement. This approach leads to better, data-driven decisions that can have an immediate, positive impact on Sun Country's marketing and operational strategies.

<Appendix>

Code

```
[ ] #Applying KMeans
kmeans = KMeans(n_clusters = 5,n_init=30)
kmeans.fit(clustering_data)

data['Cluster'] = kmeans.labels_
data.head(10000)
```

Value Counts for each cluster

```
[46] cluster_size = data['Cluster'].value_counts().sort_index()
      cluster_size
```



	count
Cluster	
0	7135
1	7481
2	4350
3	4101
4	3508

dtype: int64

Merge sample data transformed and cluster data

```
reservation_data = pd.read_csv('/content/drive/MyDrive/sample_data_transformed.csv')
```



[Show hidden output](#)

```
[ ] #Merge sample_data_transformed and cluster data
final_dataframe = reservation_data.merge(data[['uid', 'Cluster']], on='uid', how='left')
```

Example code for visualization

```
[13] #Days_pre_book by mean
      cluster_color = {
          0: 'red',
          1: 'blue',
          2: 'yellow',
          3: 'green',
          4: 'purple'
      }
      cluster_days_pre_book = data.groupby('Cluster')['days_pre_booked'].mean()

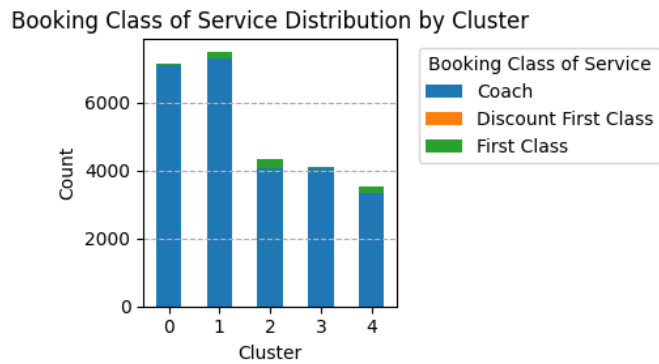
      plt.figure(figsize=(10,6))

      for cluster in cluster_days_pre_book.index:
          plt.bar(cluster,cluster_days_pre_book[cluster],color = cluster_color[cluster])

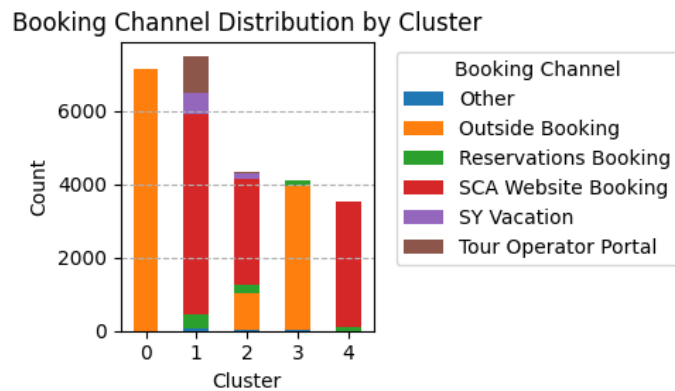
      plt.title('Number of days pre book ')
      plt.xlabel('Cluster')
      plt.ylabel('Number of days')
      plt.grid(axis='y')
      plt.show()
```

Graphs for Each Variable

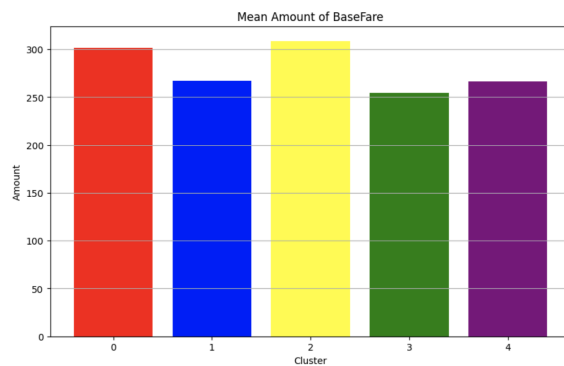
Booking Class



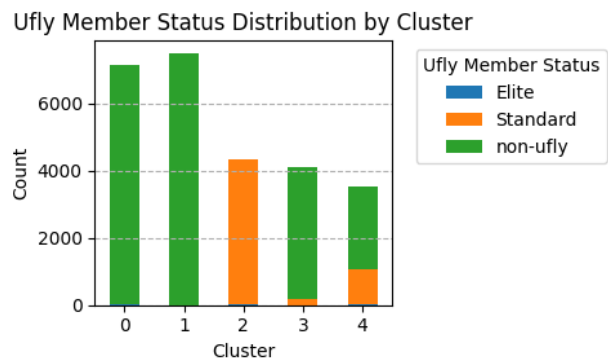
Booking Channel



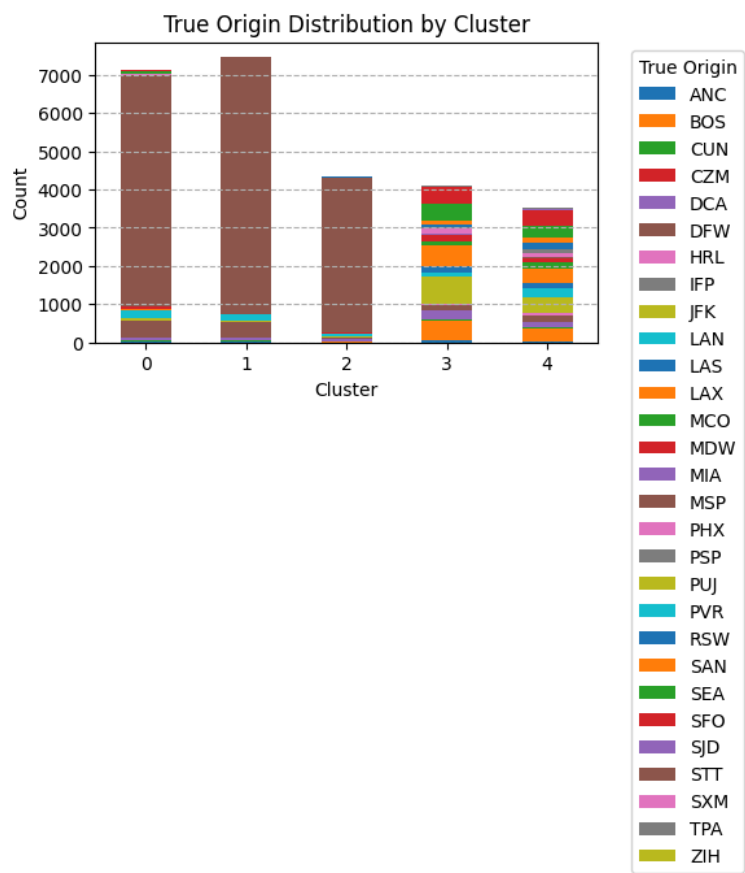
Mean Base Fare



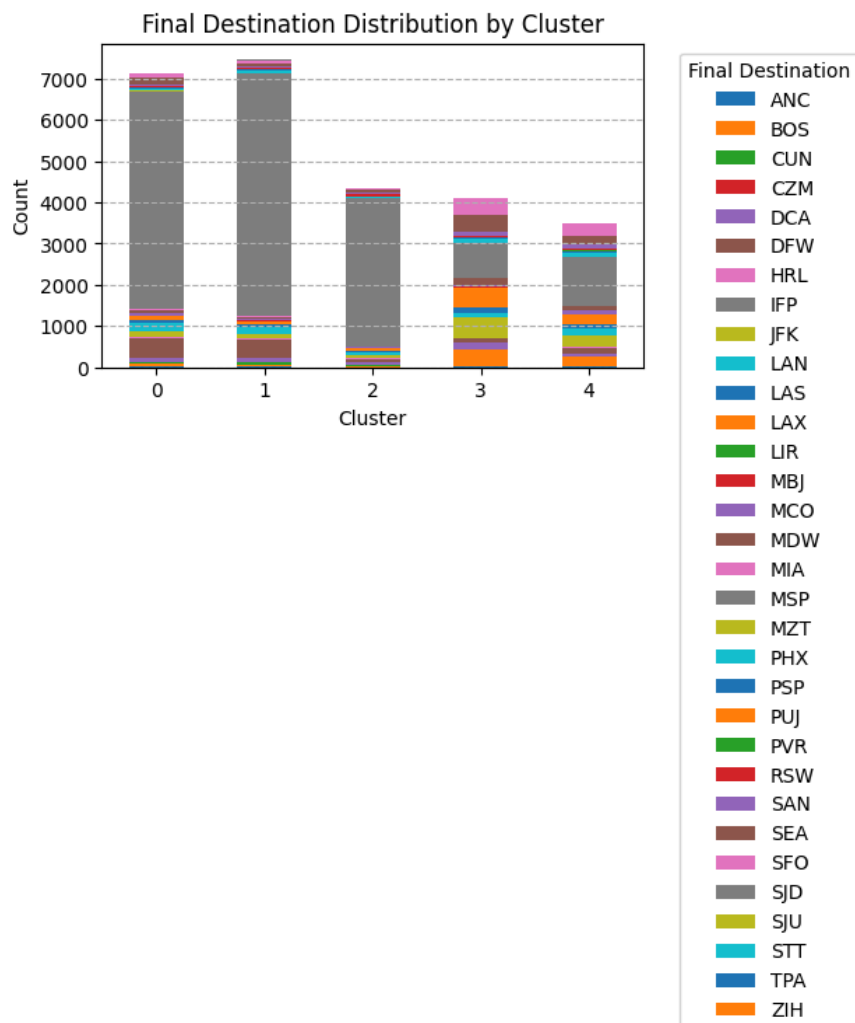
Ufly Member Status



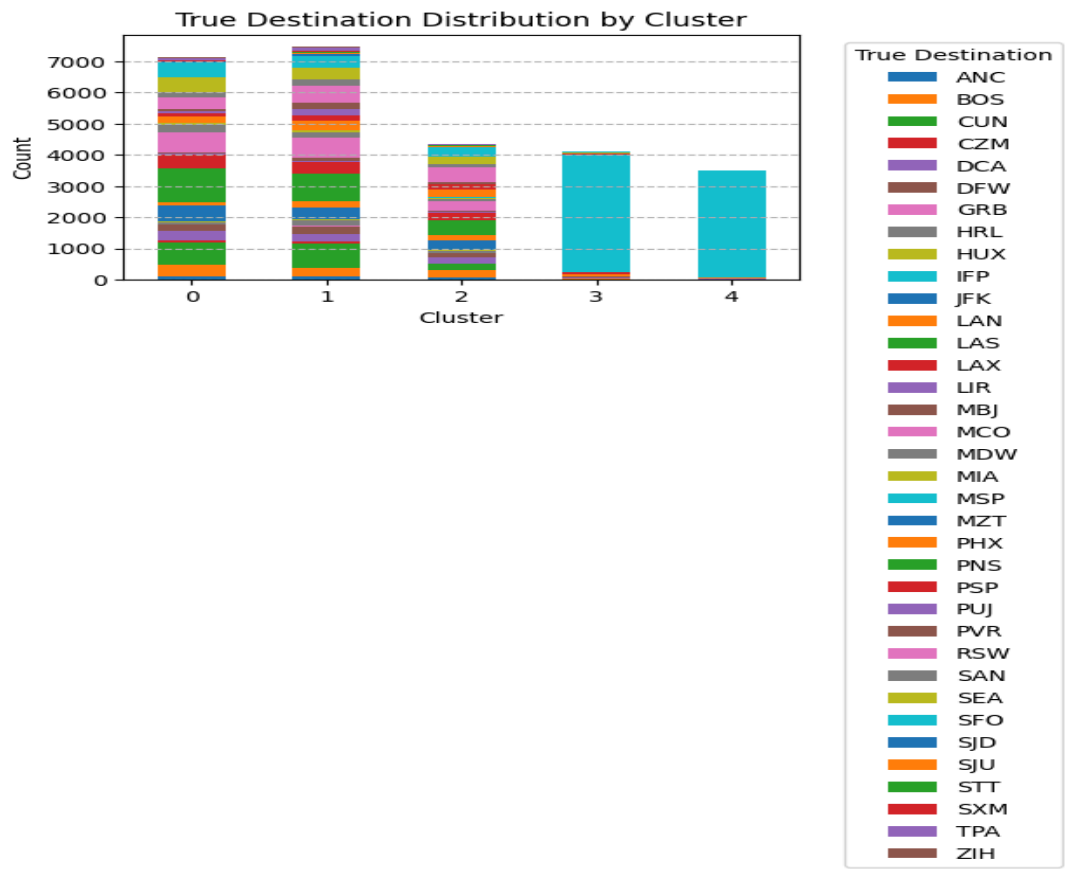
True Origin



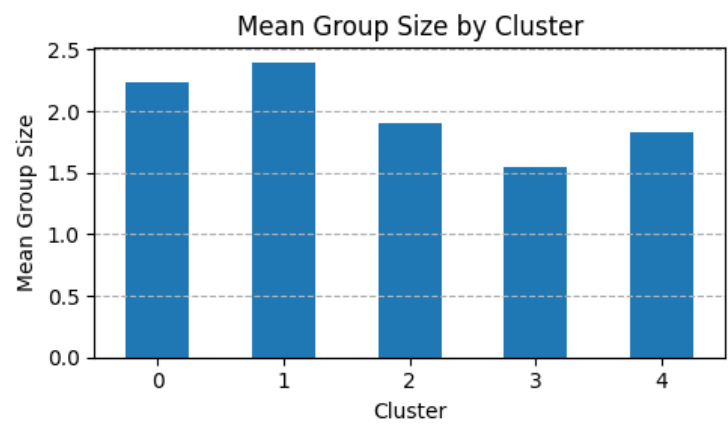
Final Destination



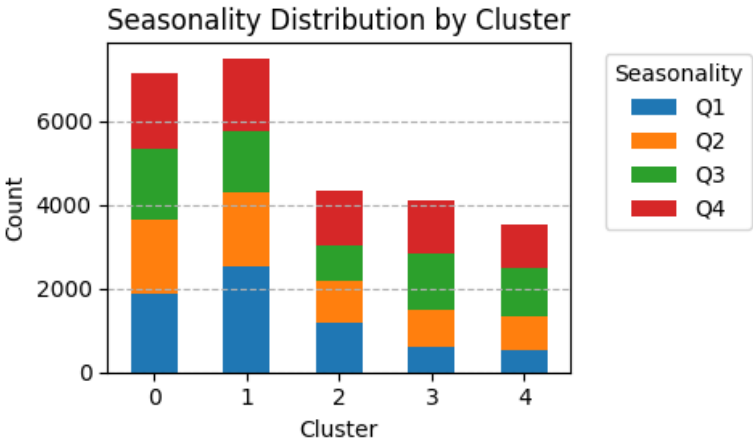
True Destination



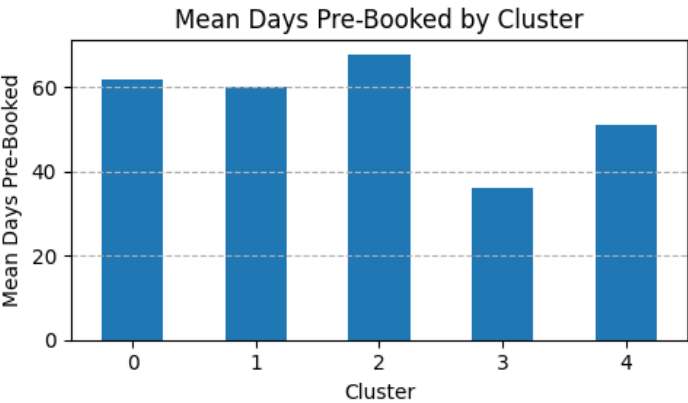
Mean Group Size



Seasonality



Mean days pre-booked



Age Group Size

