



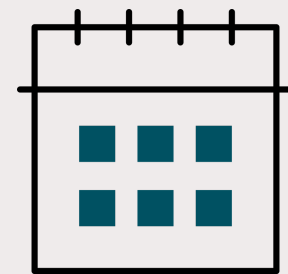
XYZ AIRLINE

UNDERSTANDING CUSTOMER BEHAVIOR & PURCHASE PATTERNS

by Addit Anwar



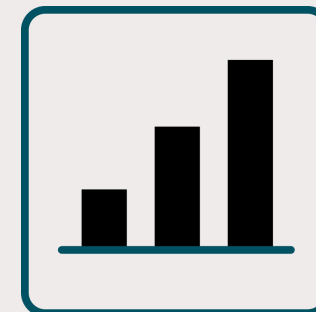
Agenda



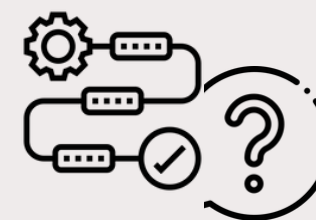
- **Business Problem Statement**



- **Data Exploration**



- **Clustering Results**

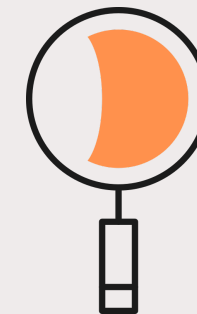


- **Recommendations**

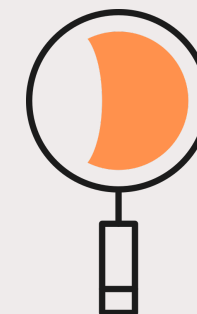
Business Problem



Assist XYZ airline by addressing lack of customer behavioral knowledge in product development



Develop a deep and reliable picture of the different segments of XYZ Airline's customers for aiding operations



Help improve customer experience by providing recommendations for membership and packages

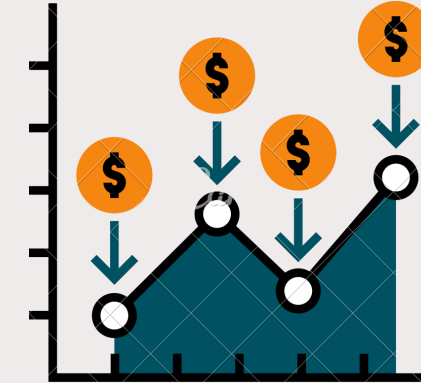
Data Overview



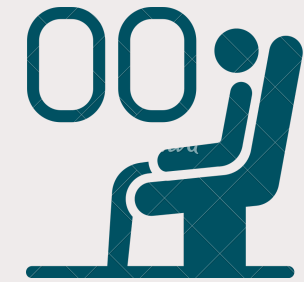
3.2M travel transactions



**Data for 2 years
(2013 & 2014)**



**Average spend:
\$155**



1.5M unique passengers



55 unique destinations

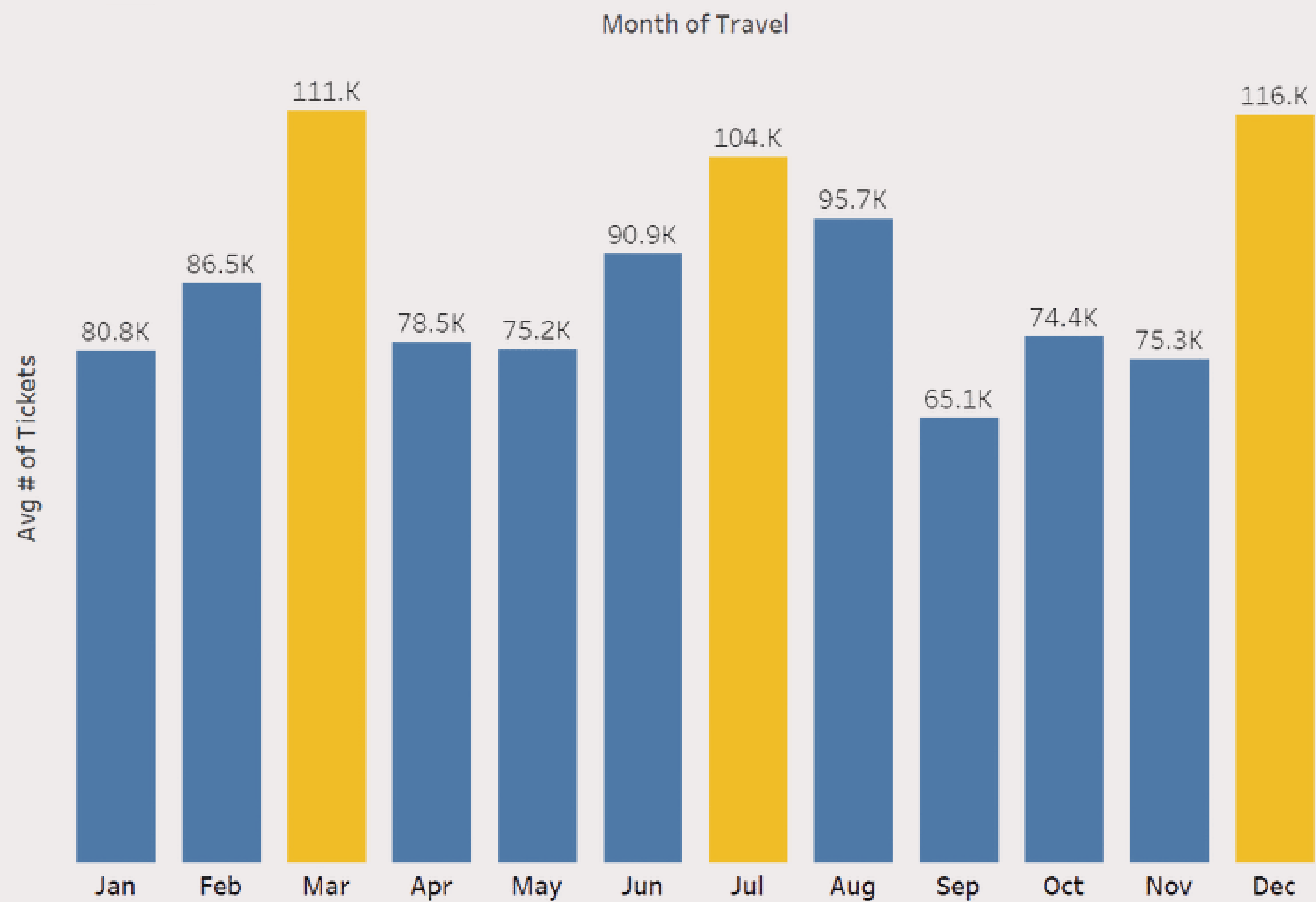


267K Premium members(17%)



**Major booking channels are:
Airports, Website(XYZ
Airline/External) and
Vacation Portals**

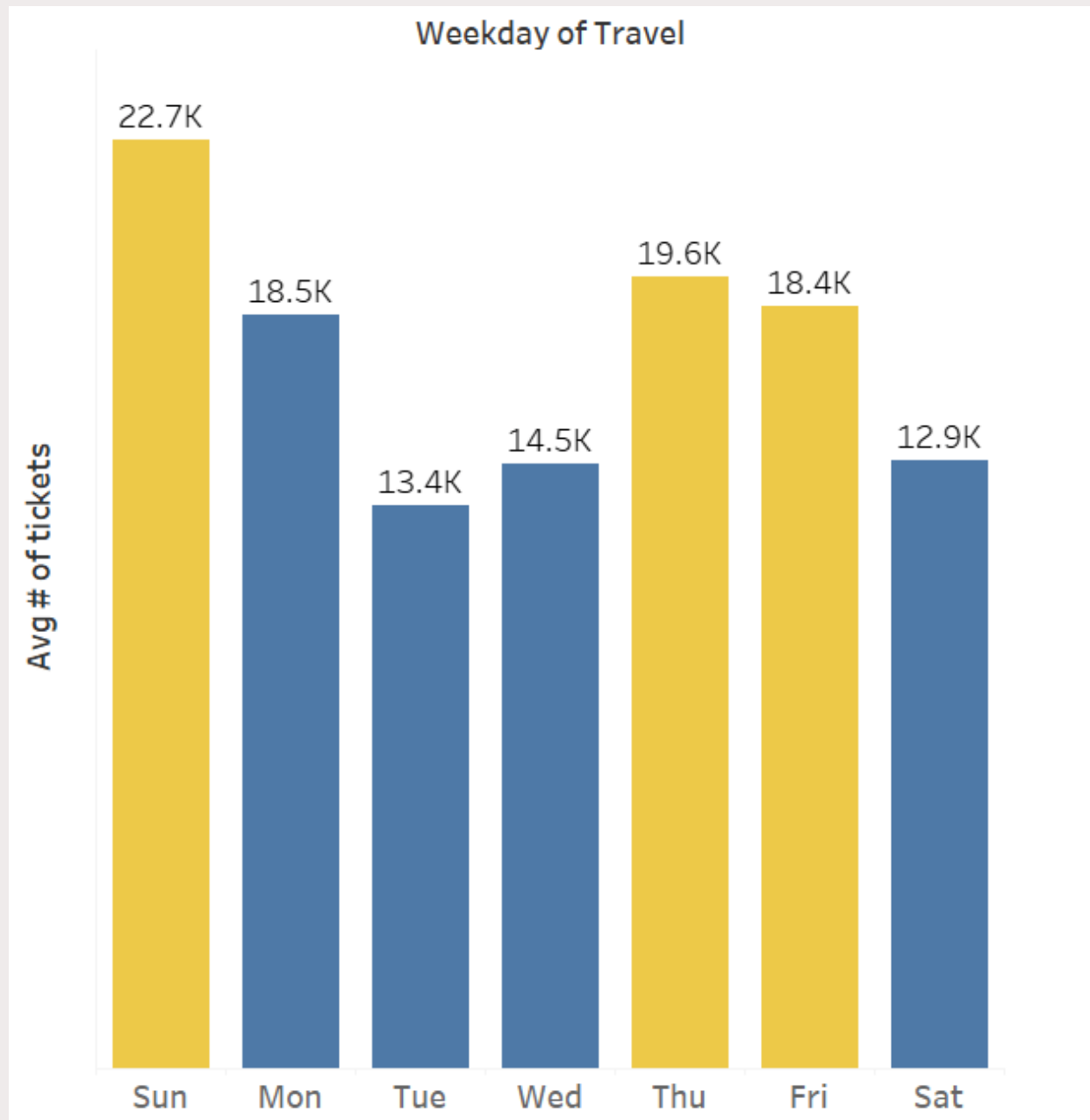
Total Travel Frequency per Calendar Month



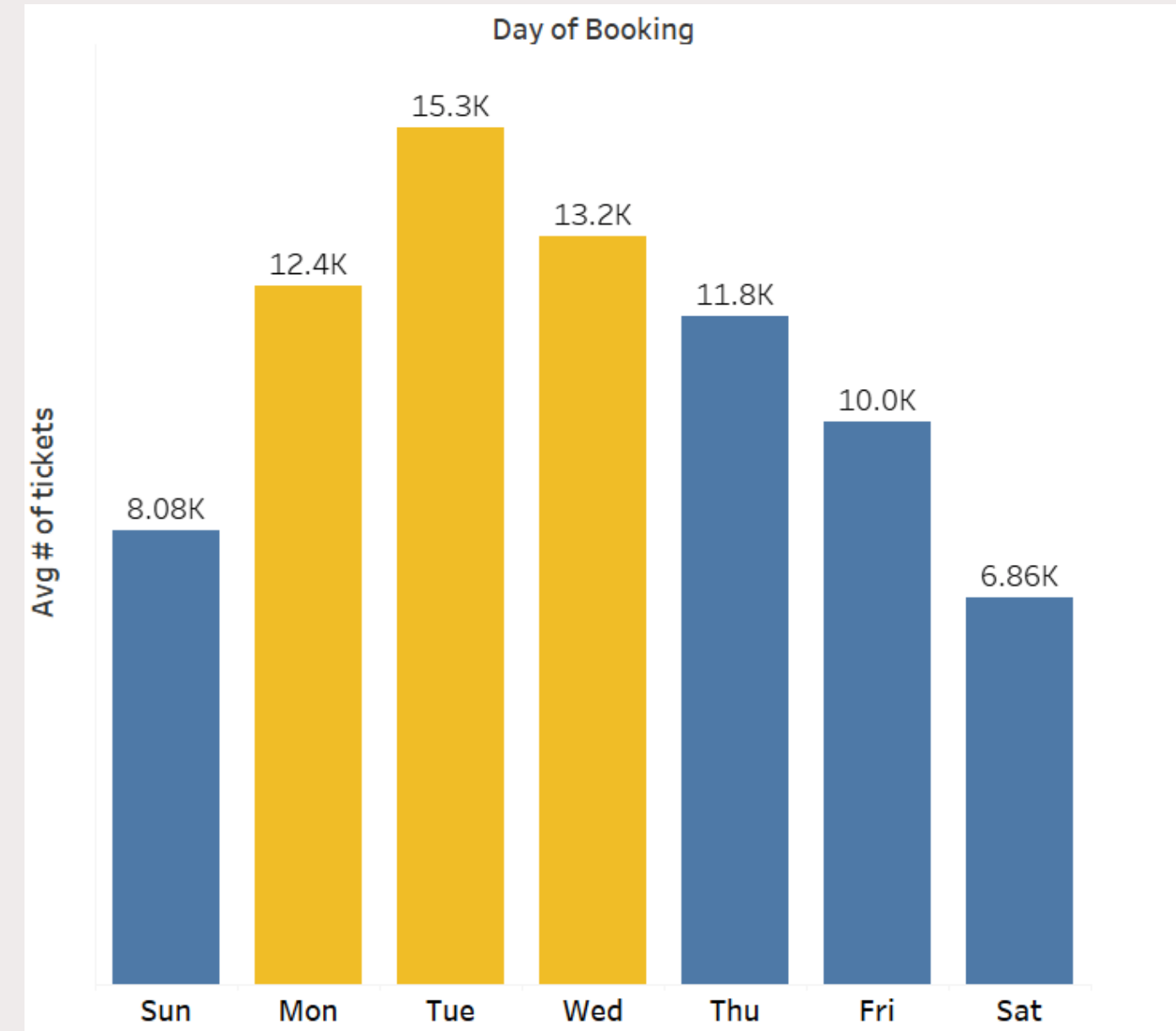
The above chart shows the Average number of tickets across year 2013 and 2014

As expected December is the most popular month for traveling (holiday season) followed by March and July

Travel Frequency & Booking day vs Day of Week

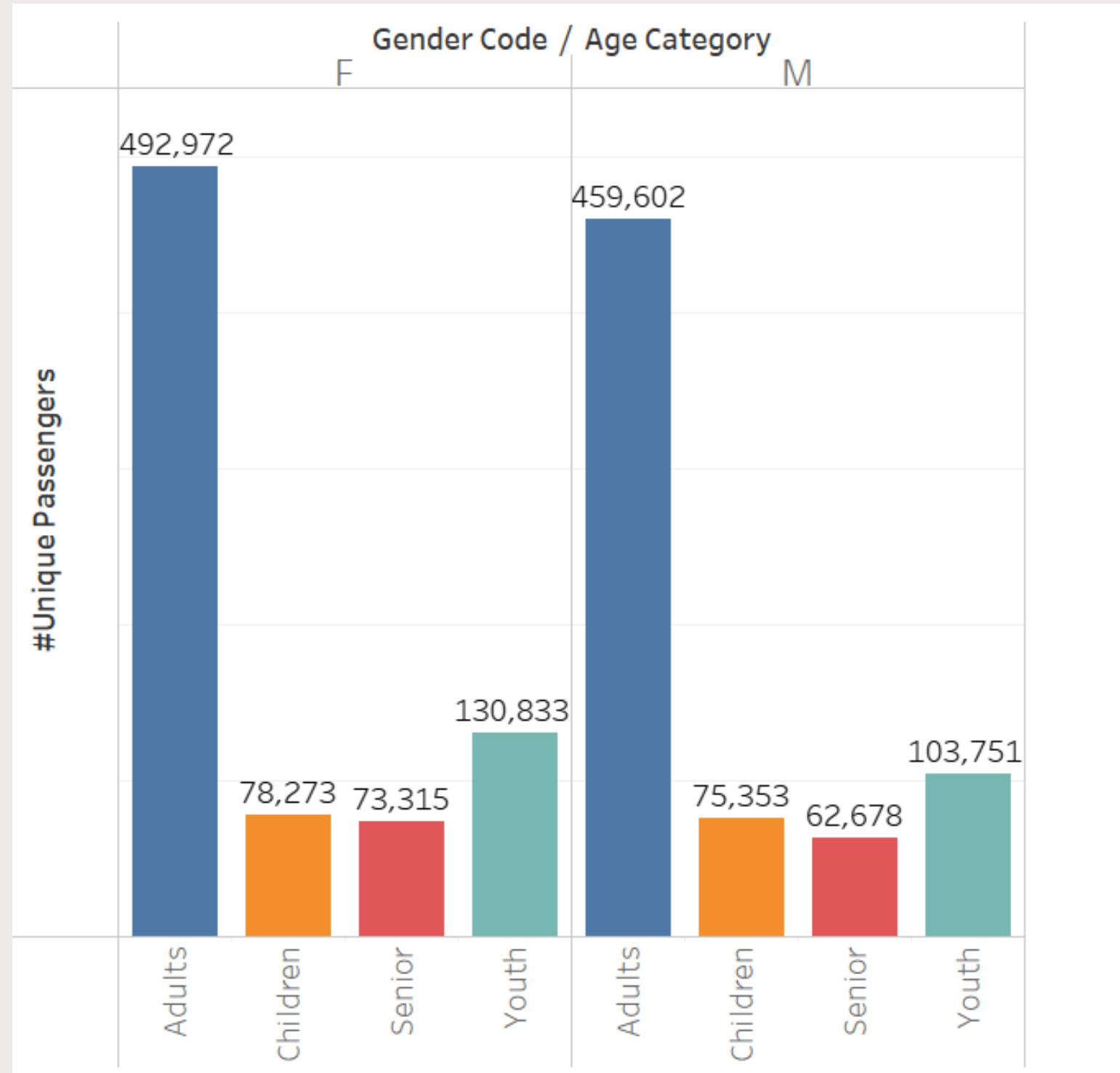


Customers travel Pre-weekend(Thurs & Fri) and return Post-Weekend (Sunday, Monday)



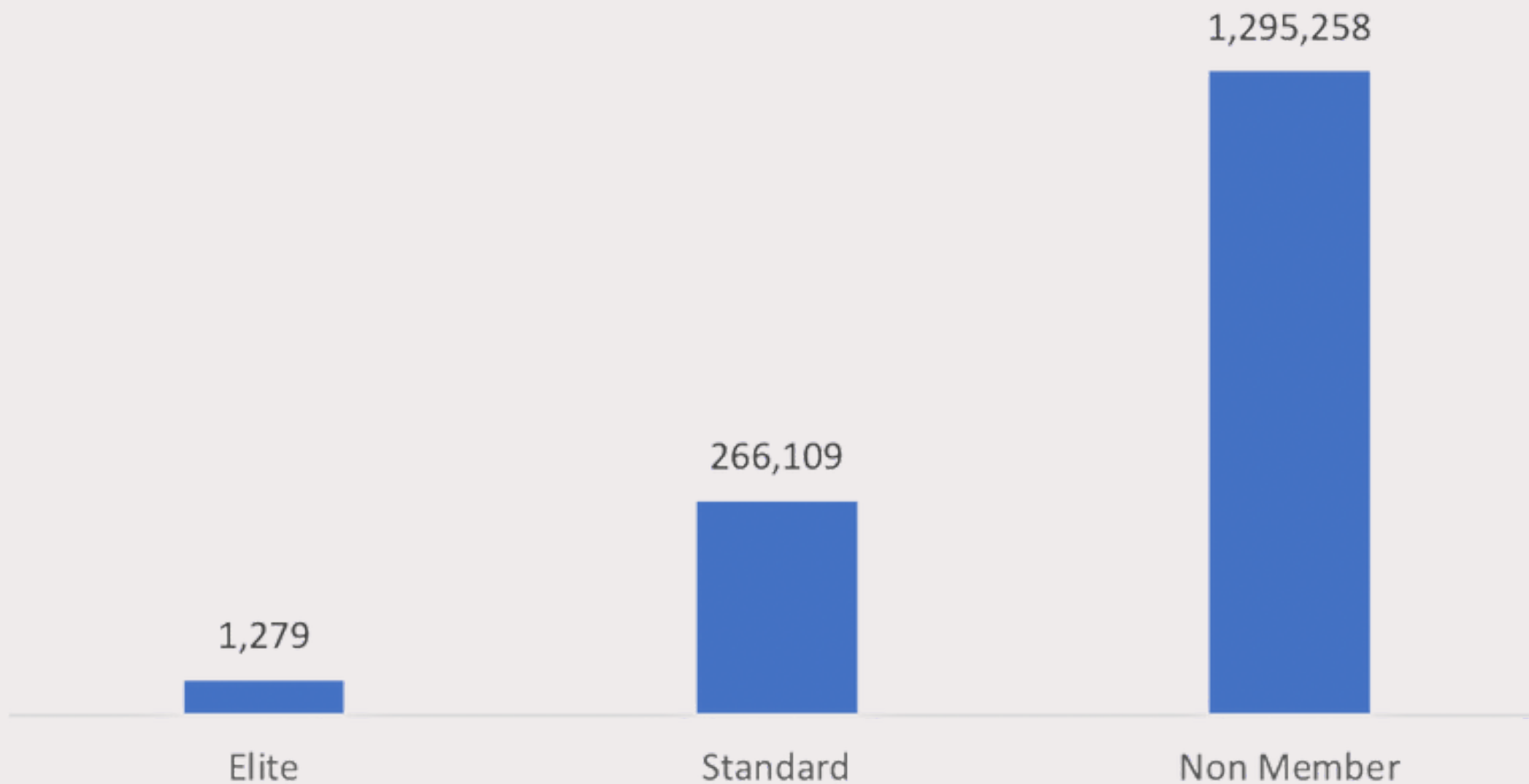
Most airline deals are announced on Monday which explains the high number of flight tickets booked on Tuesday

#Passengers vs Gender & Age



- **Age Categories:**
 - **Children** - 0 to 14 years
 - **Youth** - 15 to 25 years
 - **Adults** - 26 to 65 years
 - **Senior** - greater than 65 years
- **The number of unique female passengers is greater than the number of male passengers across all age groups**

Number of passengers across different categories



- **21% of the unique passengers are Premium members (Elite + Standard)**
- **Standard membership can be bought for an annual fee**
- **Standard members are classified as Elite members when certain travel conditions are met**

Top Serviced Routes

Most Popular Destinations By Count

Tourist

Non Tourist

LAS

SFO

CUN

JFK

MCO

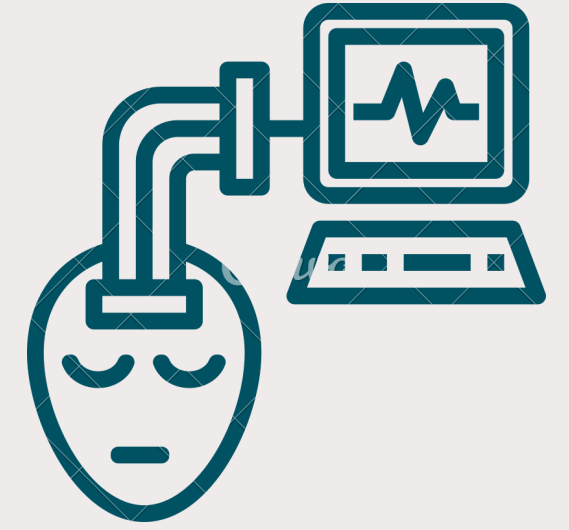
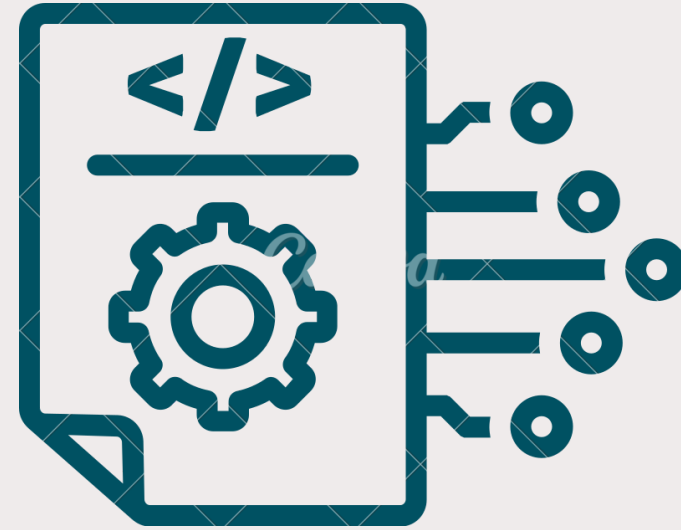
LAX

DFW

SEA

The counts for tourist are matched based on the booking channel provided in the data - SY Packages and Tourist booking

Data Modelling Process



Data Manipulation

- Split the data across membership categories.
- Performed Stratified sampling to reduce data size.

Model Optimization

- Performed Clustering analysis as it divides the data into meaningful groups
- Chose K-Prototypes for clustering due to mixed data types
- By using the elbow curve, we determined the best number of clusters (=4 for members and 6 for all passengers)

Interpretation

- Use characteristics of each cluster to derive insights

Insights & Recommendations

- Our clustering model for Members only data gave 4 clusters as output. Refer to table below.

| Cluster | Gender | Avg. Age | # Members | Avg. Days Before Booking | Avg. amount spent | Booking Channel most popular |
|---------|--------|----------|-----------|--------------------------|-------------------|------------------------------|
| 0 | Male | 45 | 12,648 | 64 | \$306 | External |
| 1 | Female | 43 | 458 | 55 | \$241 | External |
| | Male | 42 | 442 | 51 | \$251 | External |
| 2 | Female | 54 | 5,254 | 62 | \$269 | Internal |
| 3 | Female | 40 | 8,692 | 73 | \$276 | External |

- Members in cluster 0 (highlighted in yellow in Table 1) are the high spending Premium rewards members. The people in this cluster are all male and have an average age of 45.
- These members tend to book around 2 months before their travel via external channels(Outside Booking, Tour Operator Portal, Reservations Booking)

Insights & Recommendations

- Among Non-Premium members, there are around 40K male passengers between the ages of 44 and 46 who share similar characteristics to members in cluster 0. This subgroup is ~7% of the total passengers among non - members.
- Instead of targeting all the non-members for conversion to Premium Rewards members, XYZ Airline can target the non-members in the above demographic.

Vacation Packages:

- XYZ Airline passengers who take flights from Minneapolis to vacation places (LAS , CUN , MCO, DFW) also frequently travel to HRL, HUX and CZM apart from the usual destinations such as SEA(seattle), JFK(New York) and LAX(Los Angeles).
- XYZ can offer vacation packages for the people traveling to these three places as they are more inclined to travel to the popular tourist/vacation destinations.

Insights & Recommendations

| Cluster | Gender | Avg. Age | # Members | Avg. Days Before Booking | Avg. amount spent | Booking Channel most popular |
|---------|--------|----------|-----------|--------------------------|-------------------|------------------------------|
| 0 | Male | 45 | 12,648 | 64 | \$306 | External |
| 1 | Female | 43 | 458 | 55 | \$241 | External |
| | Male | 42 | 442 | 51 | \$251 | External |
| 2 | Female | 54 | 5,254 | 62 | \$269 | Internal |
| 3 | Female | 40 | 8,692 | 73 | \$276 | External |

- We also created clusters on all passengers and found a few interesting clusters :
 - Cluster 0 - In this cluster we have mostly young females - with an average age of 22. They travel as much as Members but do not seem to be taking membership. We also noticed from the distribution of members - that the average age of cluster members is around 45. This suggests that younger people do not frequently take up membership. We would suggest XYZ Airline to market membership and its benefits to them so that we can increase the total number of this cluster's members.
 - Cluster 2 - In this cluster we have male - non members around the age of 31 who choose to book in internal channels and also get the advantage of partial discounts.Despite the partial discounts, we observe higher average amount than usual.

THANK YOU

