UNITED

TEAM SKY_AVIATORS

TO THE MOON !!!

TEAM MEMBER 1:
TUSHAR KAUSHIK

TEAM MEMBER 2:
RACHIT RASTOGI

Problem Statement

Optimize United Airlines' Call Center Performance

At United Airlines, our call center serves as the vital link between passengers and the assistance they require. However, as we strive for excellence, we encounter significant challenges that hinder our mission. Picture a passenger anxiously waiting on hold for help with a flight change. As the **Average Speed to Answer** (**AST**) drags on, their frustration mounts, eroding trust in our service.

When they finally connect with an agent, the **Average Handle Time** (**AHT**) continues to increase as complex issues are navigated, resulting in longer calls and decreased satisfaction. This scenario reflects a broader issue within our operations, where **lengthy call durations** impact both *efficiency* and the *customer experience*.

Compounding these challenges is the misallocation of resources. Many inquiries that could be resolved through self-service options escalate unnecessarily to agents, overwhelming our team during peak times. To elevate our call center operations and ensure every traveler feels valued, we must tackle these issues of **long wait times**, **extended call durations**, and **ineffective routing** of inquiries head-on.

AHT and AST Analysis

Identify key drivers of long Average Handle Time (AHT) and Average Speed to Answer (AST).

Self-Service Optimization

Propose IVR improvements to reduce agent intervention for self-solvable issues.

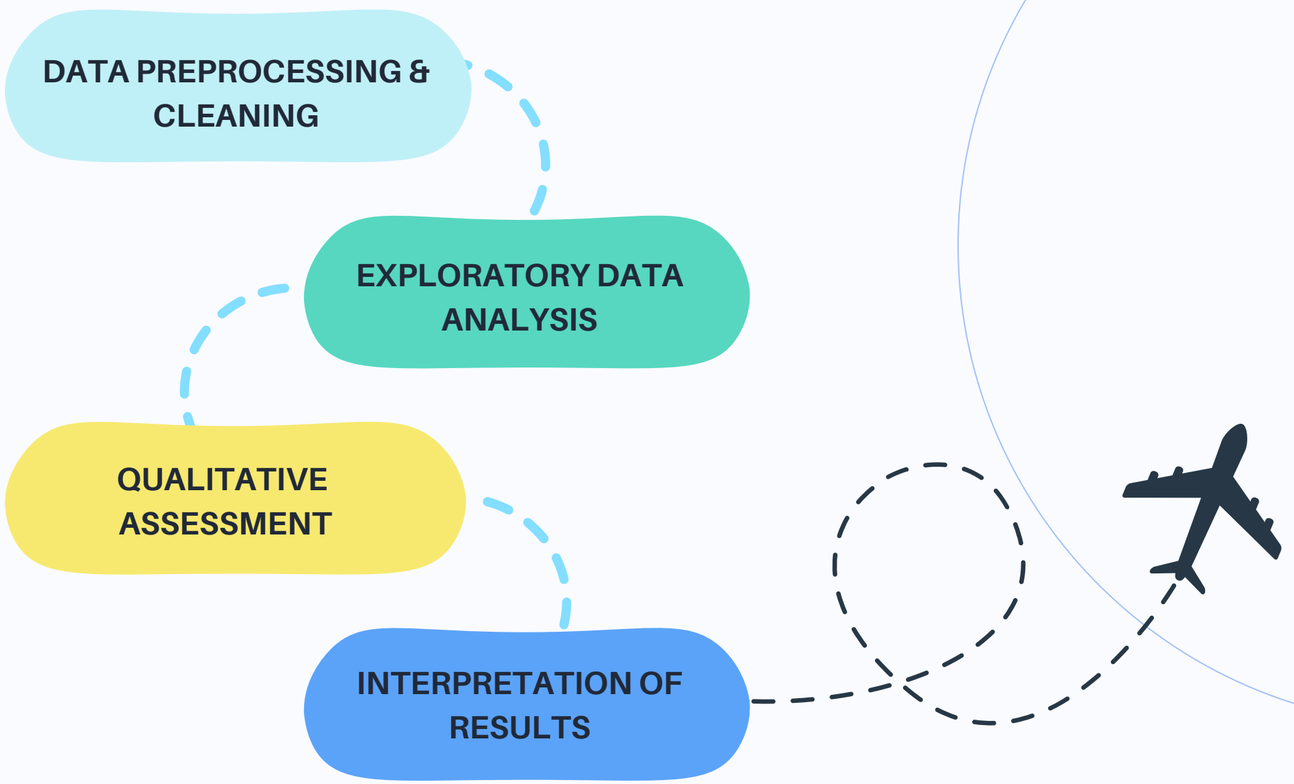
Call Reasons Categorization

Analyze call transcripts to uncover and categorize primary call reasons for operational efficiency.

Our Approach

Recommendations Based On Data Exploration

Here is the workflow diagram for our process we came up with:



NLTK

pandas

DATA PREPROCESSING & CLEANING

Cleaning, Preprocessing & Preparation

During data preprocessing and cleaning , we began by importing the pandas library and loading our dataset using **pd.read_csv**(). With the data in hand, we carefully examined its structure and types, utilizing **df.info**() to uncover insights about its contents and **df.head**() to grasp the context.

As we delved deeper, we addressed various data quality issues: handling missing values through thoughtful imputation or deletion, standardizing data formats for consistency, removing duplicate entries to maintain integrity, and identifying outliers that could skew our analysis. Recognizing the importance of a comprehensive dataset, we merged relevant datasets using **pd.merge**() based on common keys, ensuring that our combined data retained its integrity.

Finally, we conducted a thorough re-examination of our cleaned dataset, once more to confirm it was primed and ready for the next phase of our analysis.

```
result_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 66455 entries, 0 to 66454
Data columns (total 15 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   call_id                 66455 non-null  int64  
 1   customer_id             66455 non-null  int64  
 2   agent_id                66455 non-null  int64  
 3   call_transcript         66455 non-null  object 
 4   call_hour               66455 non-null  int32  
 5   waiting_time            66455 non-null  float64
 6   handling_time           66455 non-null  float64
 7   primary_call_reason     66455 non-null  object 
 8   agent_tone              66455 non-null  object 
 9   customer_tone           66455 non-null  object 
 10  average_sentiment       66455 non-null  float64
 11  silence_percent_average 66455 non-null  float64
 12  elite_level_code        42579 non-null  float64
 13  elite_level_category    66455 non-null  object 
 14  preprocessed_transcript 66455 non-null  object 
dtypes: float64(5), int32(1), int64(3), object(6)
memory usage: 7.4+ MB
```

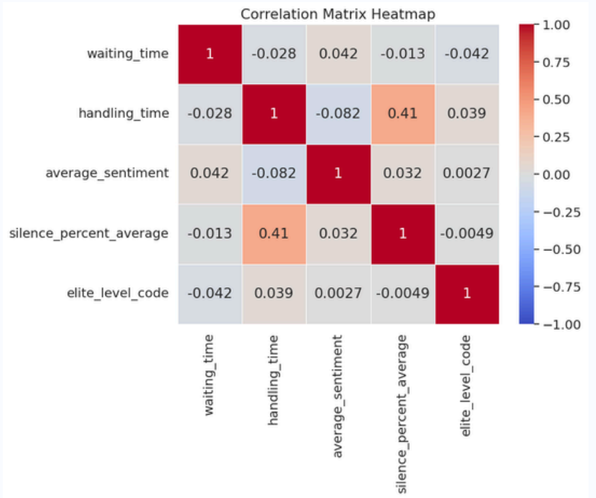
EXPLORATORY DATA ANALYSIS

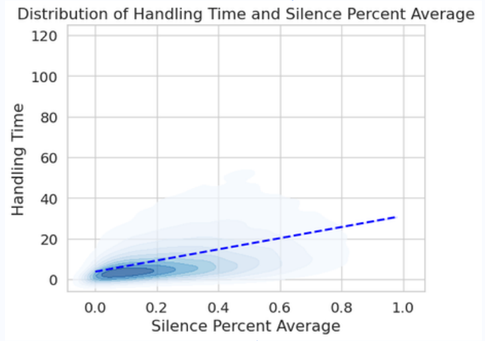
Visualisating Key Parameters

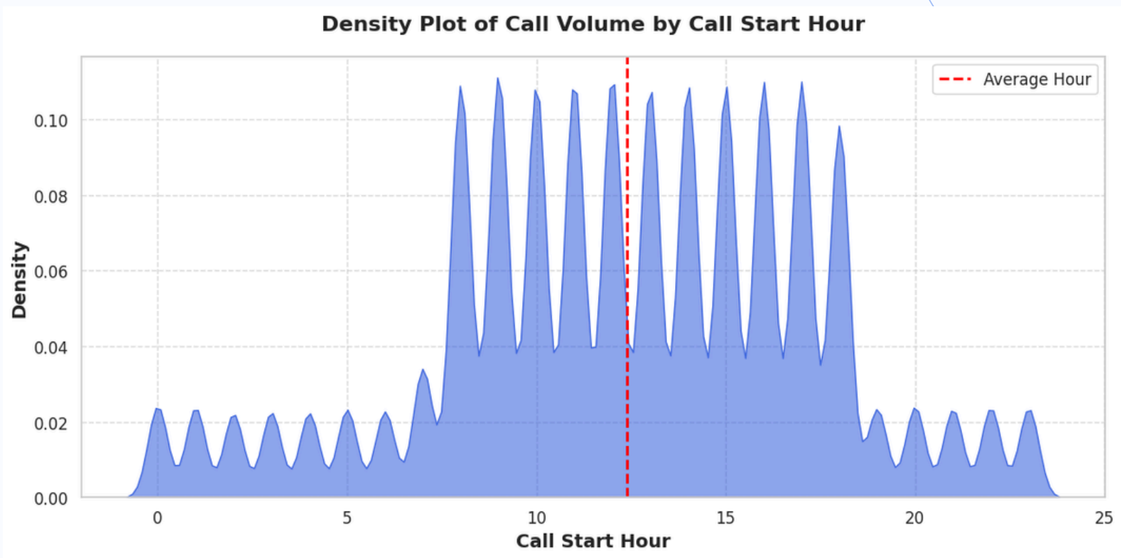
We visualized key metrics to uncover insights for improving customer service. Starting with histograms of waiting and handling times, we identified patterns revealing process bottlenecks. Scatterplots examined relationships between handling time and silence percentage, showing how **silence effects AHT**. Density plots, pie charts, & heatmaps highlighted **call duration volume**, primary call reasons, & **correlations**.

After a thorough analysis we noticed that there was no visible key driver affecting **AST** directly.

We analyzed **top call reasons by frequency** and **sentiment**, with bar plots showing the **impact of customer and agent tone on handling time.** These visualizations provided a comprehensive understanding of our operations, enabling data-driven improvements.





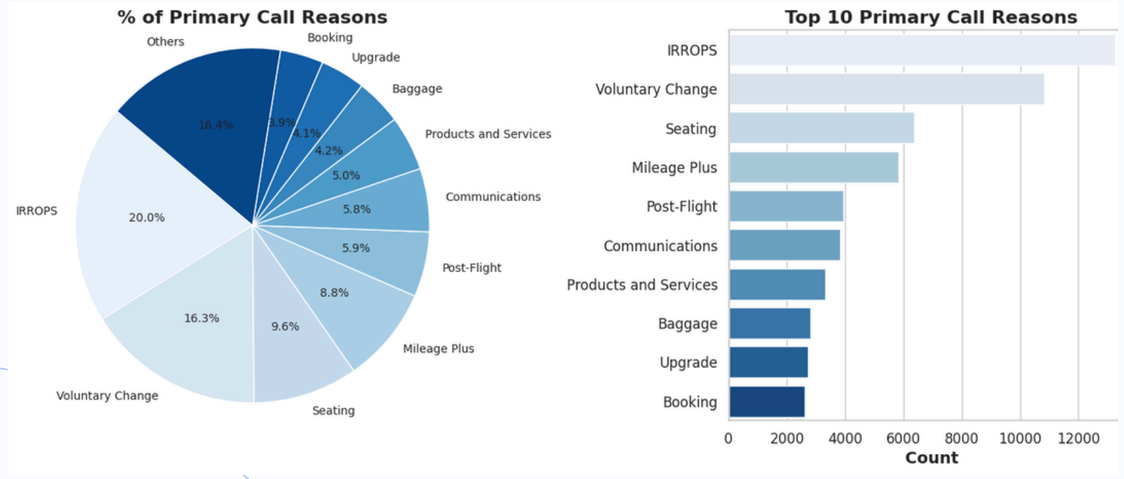


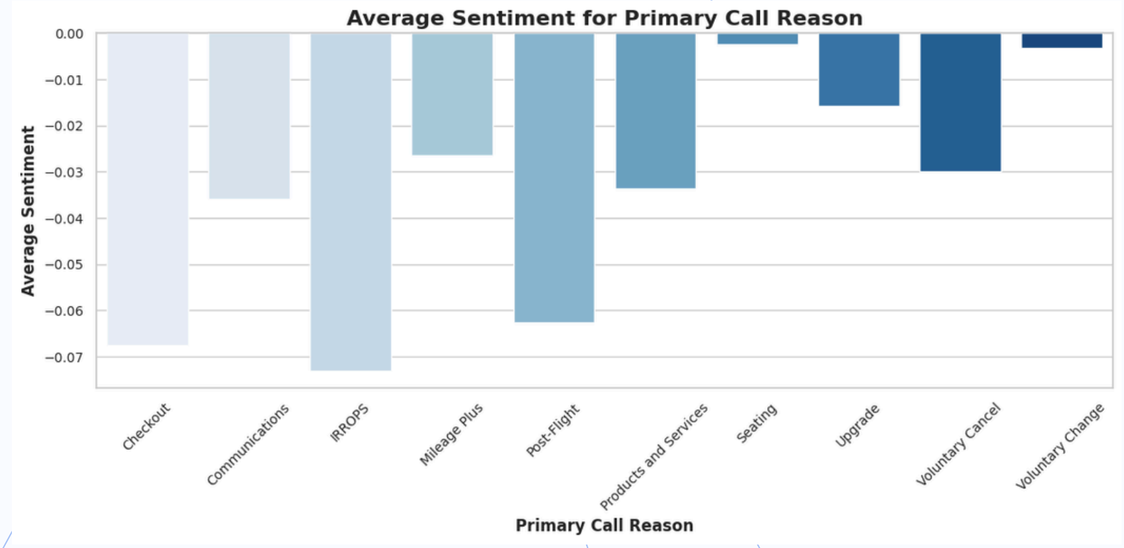
Primary Call Reasons Affecting Customers

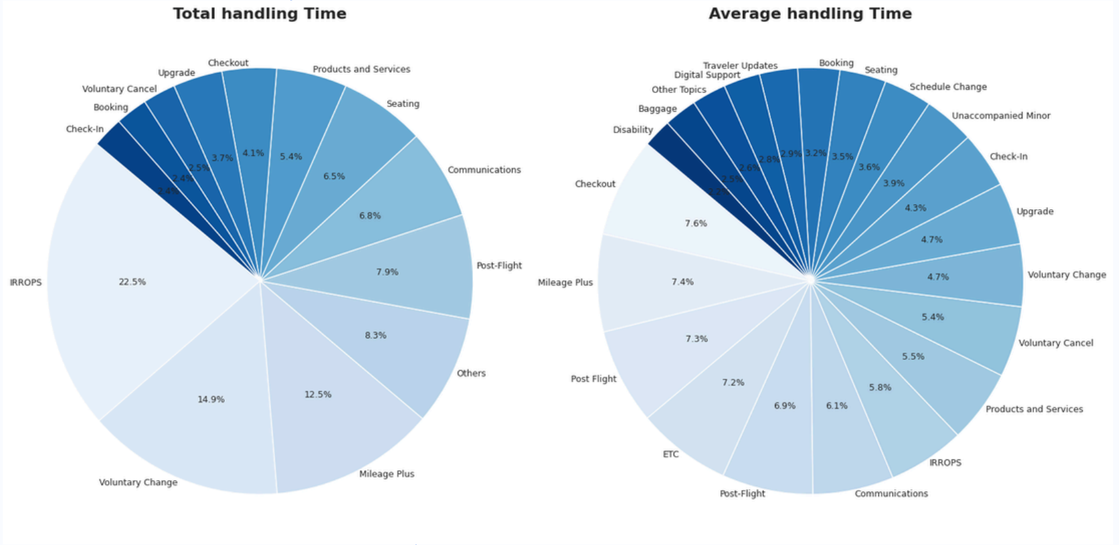
Identifying Key Call Reasons which hinders our excellent Services provided to our customers.

Prioritizing & **Resolving** Primary call reasons will **Enhance** our services by **Reducing** Total Handling time.

Primary Call Reason directly affects Customer Tone & Sentiment. **Tackling** top Primary Call Reasons will **improve Customer Satisfaction**.







Understanding Primary Call Reasons

Analyzing the variations between **Customer tones** & **Elite level Code** for each Primary Call Reason

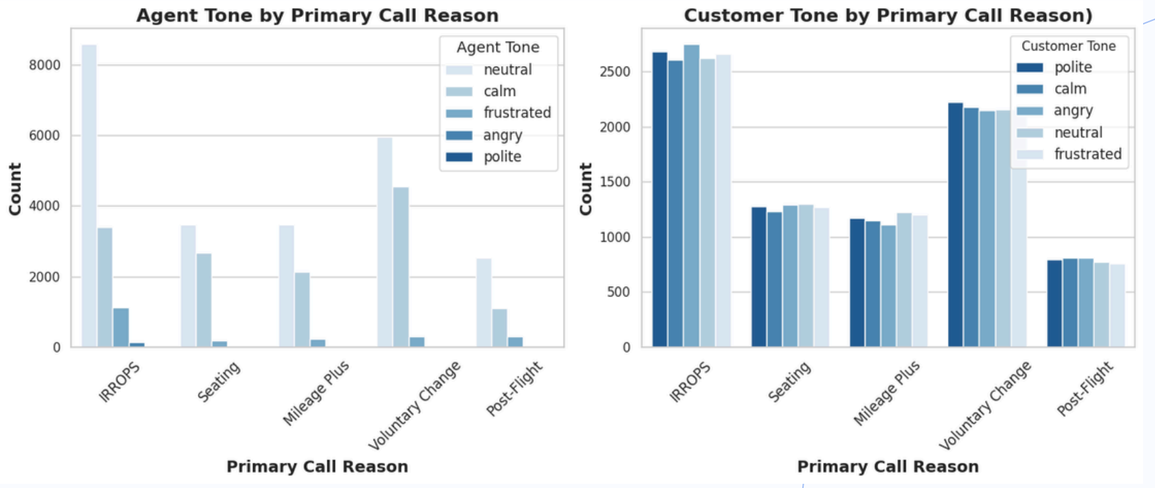
- Our findings state, **Different customers** within respective **Elite level code** has **different priorities.**

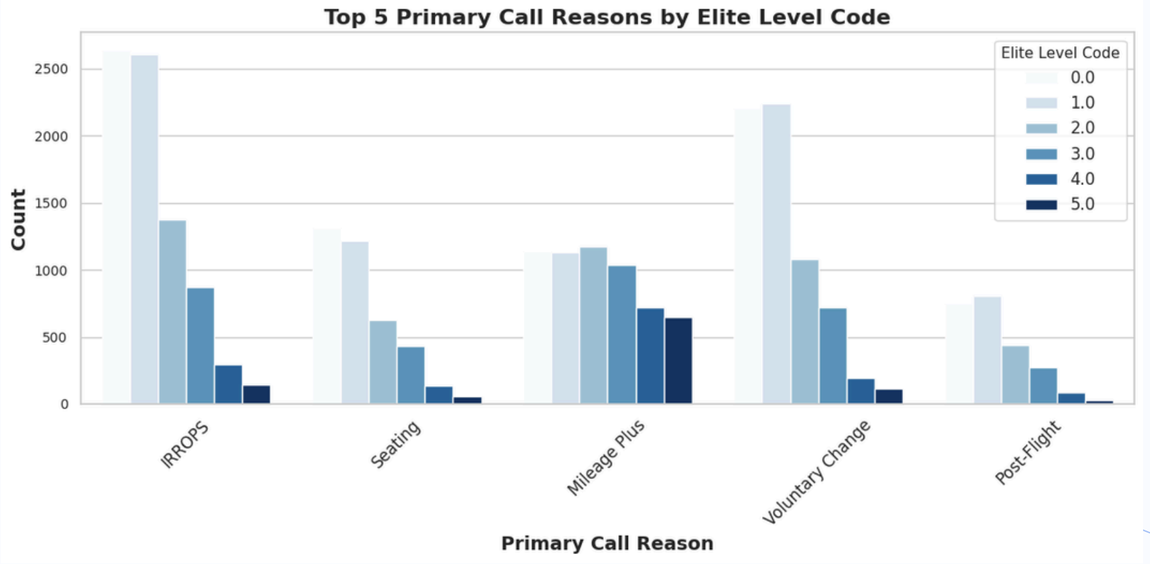
- **IRROPS** & **Voluntary Change** is a recurring issue within customers with **low Elite level code**. Whereas, **Mileage Plus** is disproportionately significant issue within **high Elite level code.**

- **Agent tone** is mostly neutral while addressing Primary Call reason. IVR **systems** should incorporate **Polite** and **Calm** tone while handling customers.

Most Frequent Reason: **IRROPS** with AHT: **13.11** minutes
Least Frequent Reason: **Unaccompanied Minor** with AHT: **8.65** minutes
Percentage difference between the average handling time for the most & least frequent call reason : **40.95%**

The **sentiment** of the conversation is almost solely **dependent** on **Agent tone**. Let's see how...

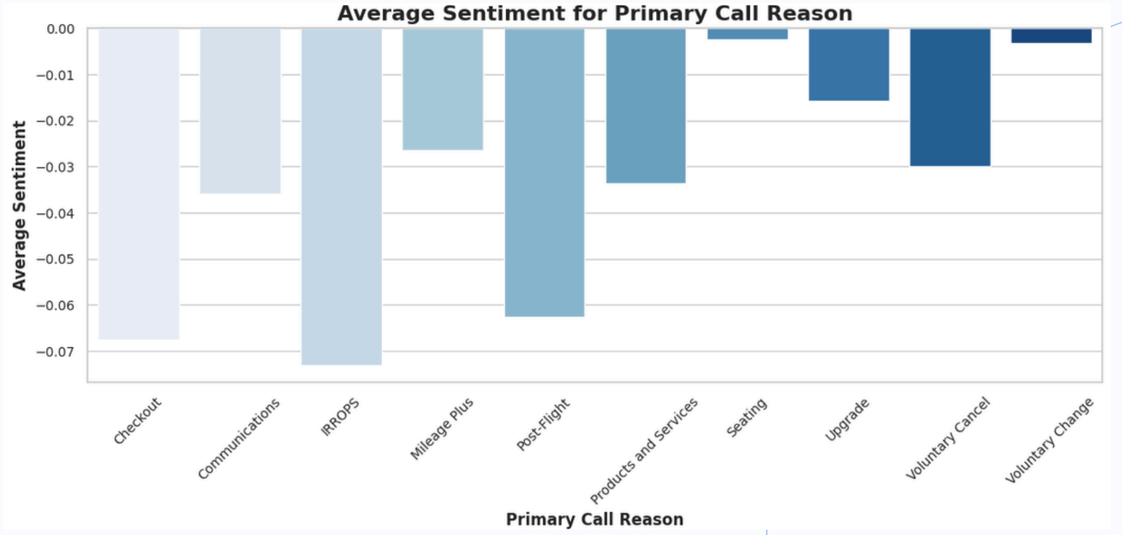


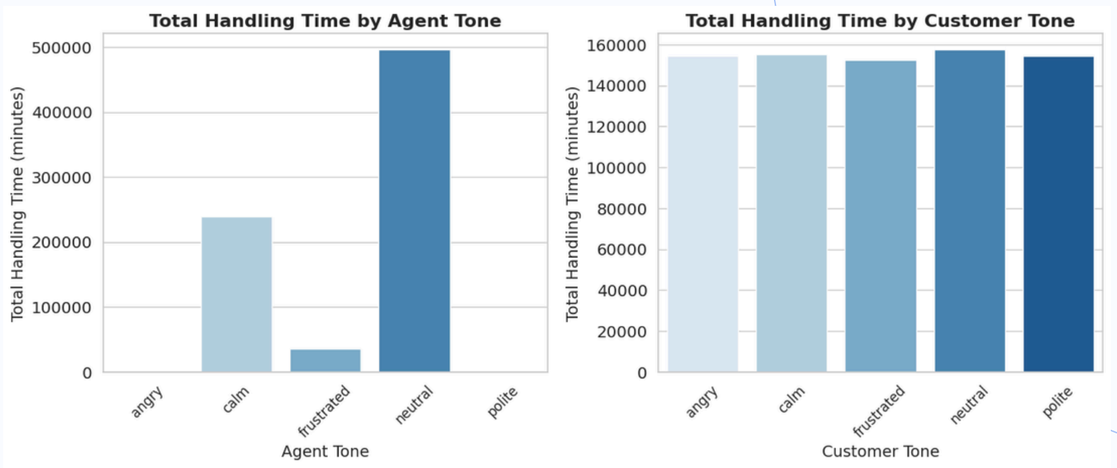


Understanding Factors affecting Sentiments

Realizing how **Primary call reasons** and **sentiment** are correlated

- Agent Tone can be classified into - **Calm**, **Frustrated** & **Neutral**. Almost no polite tones, **Increase** in **Polite** & **Calm** tones, self-solvable issues can be greatly increases.
- There is almost **equal** amounts of **customers** speaking in different **tones.**
- Therefore, to **enhance Sentiment** for Primary call reason such as **IRROPS, Checkout** & **Post-Flight** issues, agent tone should be **Polite** & **Calm**.



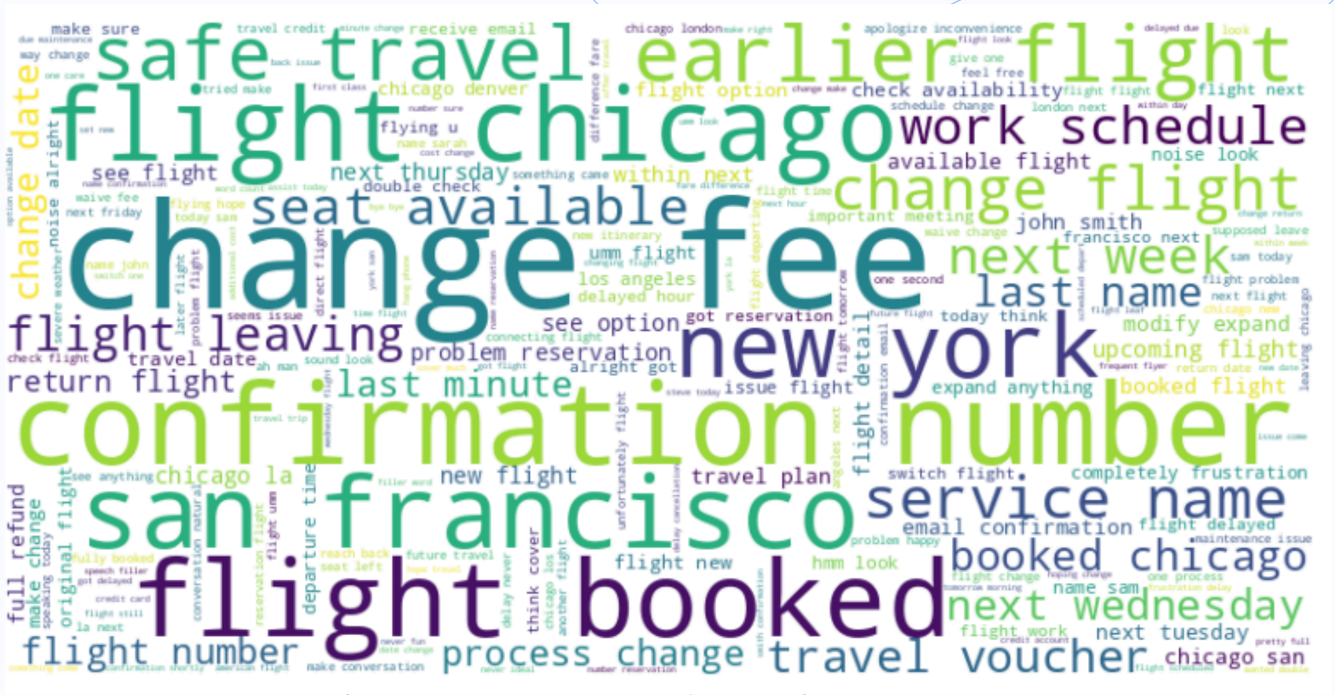


Qualitative Assessment

After a rigorous analysis of all call transcripts, there are some issues which are repeating. **Focusing** & choosing to **solve** these issues will be in direct accordance to **enhanced Customer satisfaction**, **Operational efficiency** & **number of bookings**!
This will automatically **reduce agent intervention** & **dependency on IVR systems**.

Interpretation of Result

After a rigorous analysis of all call transcripts, there are some issues which are repeating. Focusing & choosing to solve these issues will be in direct accordance to enhanced Customer satisfaction, Operational efficiency & number of bookings! This will automatically reduce agent intervention & dependency on IVR systems.



Thank You So Much !!