SKYHACK 2.0

THEME: *Call Center Optimization*

Vaibhav Gupta
(vaibhavgupta_ep21b11_50@dtu.ac.in)

Kshitij Agarwal
(kshitijagarwal_ep20a18_52@dtu.ac.in)

Repository link: GitHub

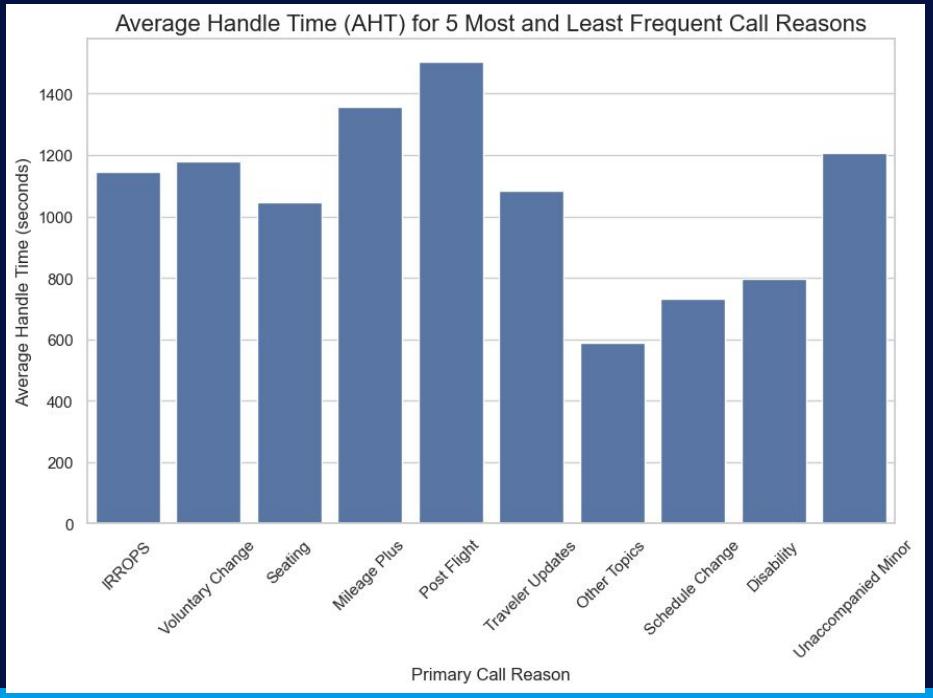
Executive Summary

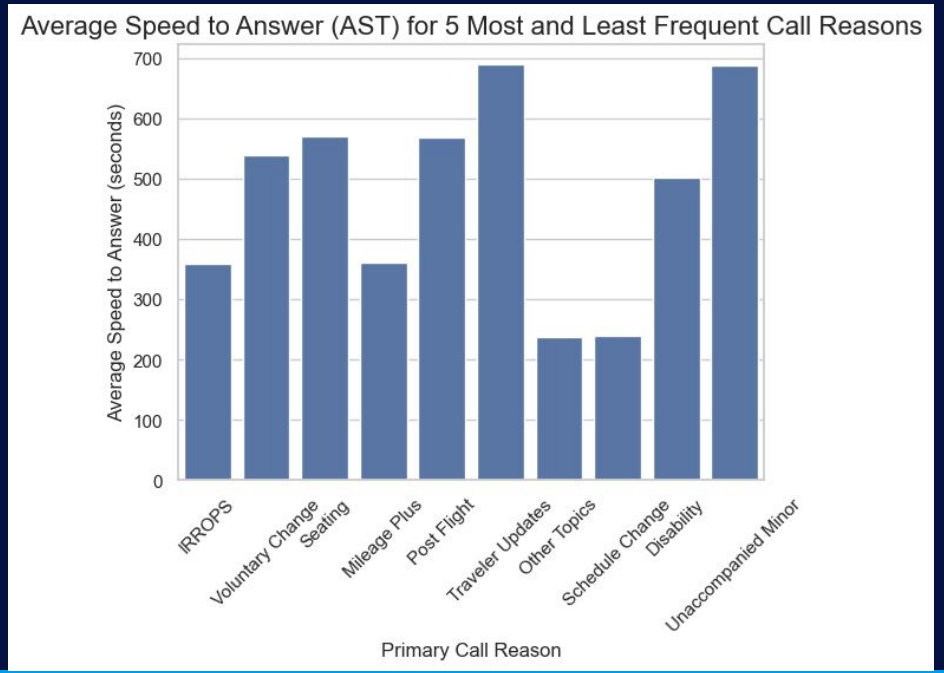
1. **93.24%** of total calls could potentially be resolved via IVR.
2. Factors Contributing to Extended Average Handle Time (AHT) include **High Call Volumes** and **Negative Average Sentiment**.
3. Recommendations for Optimizing Call Center Performance include Enhance IVR System, Proactive Customer Communications and Peak Hour Optimization.
4. Predicted 5,157 Missing Call Reasons: Used TF-IDF for text vectorization and Random Forest for classification.

Key Factors Contributing to Extended Average Handle Time (AHT):

1. **High Call Volumes**: Calls with most frequent call reasons have longer handling times (AHT).

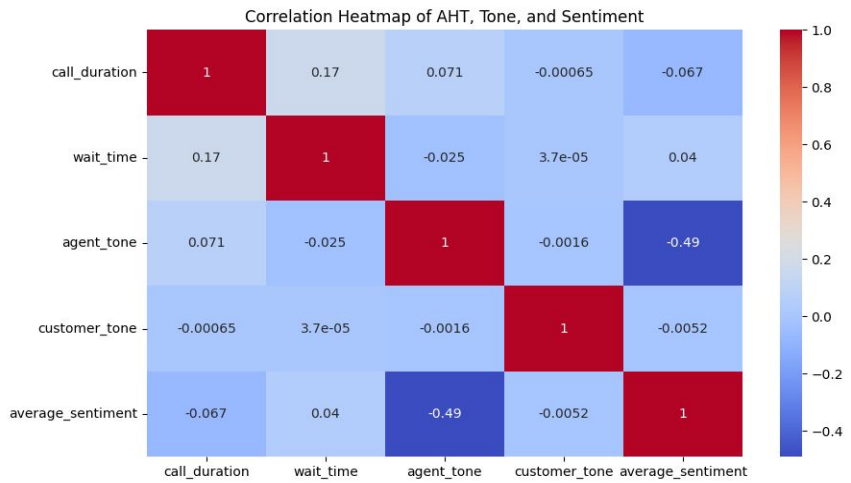
```
Average AHT for the 5 most frequent call reasons: 1245.91 seconds
Average AHT for the 5 least frequent call reasons: 881.22 seconds
Percentage difference between the average AHTs: 41.39%
```

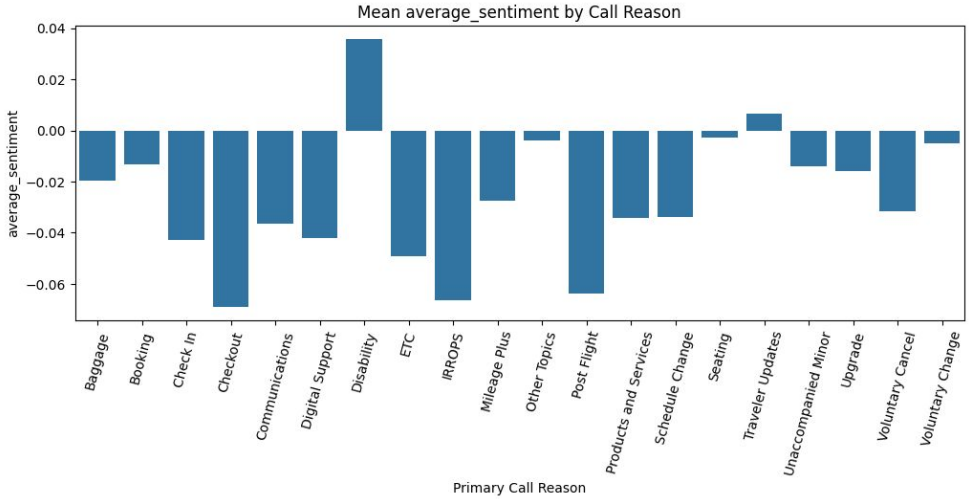


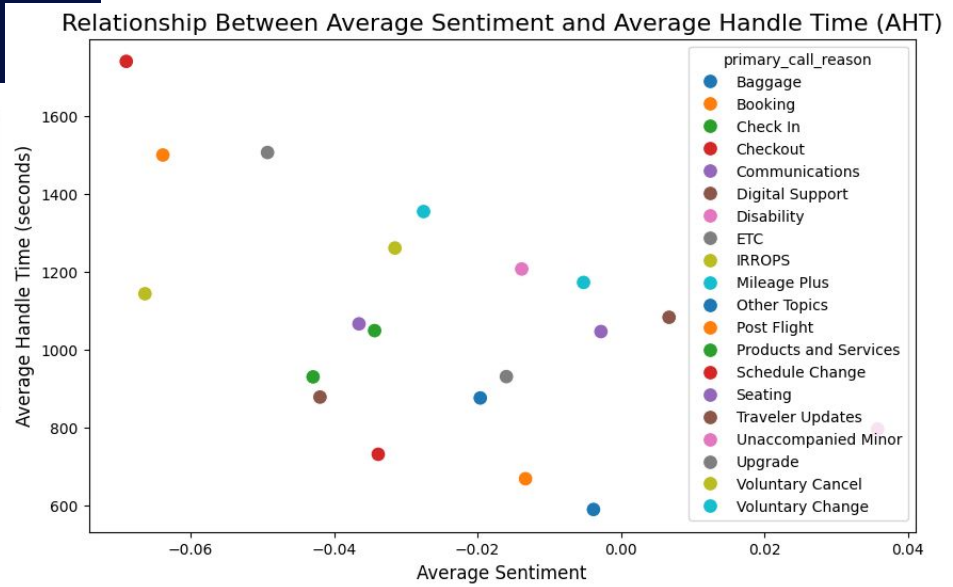


2. Negative Average Sentiment: Calls with Average Negative Sentiment have longer handling times (AHT).

- Correlation matrix shows agent tone greatly affects average sentiment of the call.

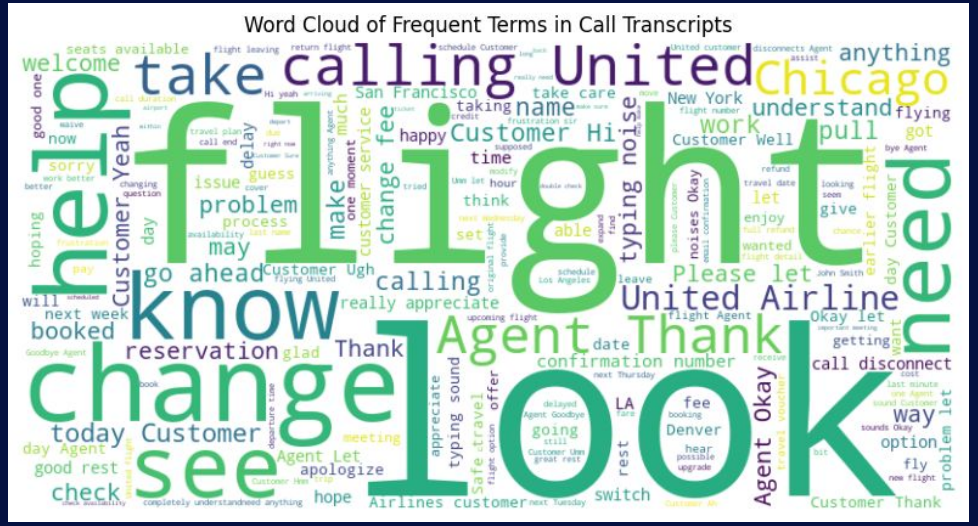






Key Findings for self solvable issues

- **Top Call Reasons by Volume**: IRROPS, Seating, and Mileage Plus accounted for a significant portion of calls.
- **Common Themes in Call Transcripts**: Terms like "flight", "agent", and "refund" appeared frequently across clusters, indicating frequent issues with flight changes and refunds.
- **Clustering Results**: Revealed key categories (flight changes, delays, baggage issues) that could be optimized through self-service.
- **93.24% of Calls Could be Solved via IVR.**

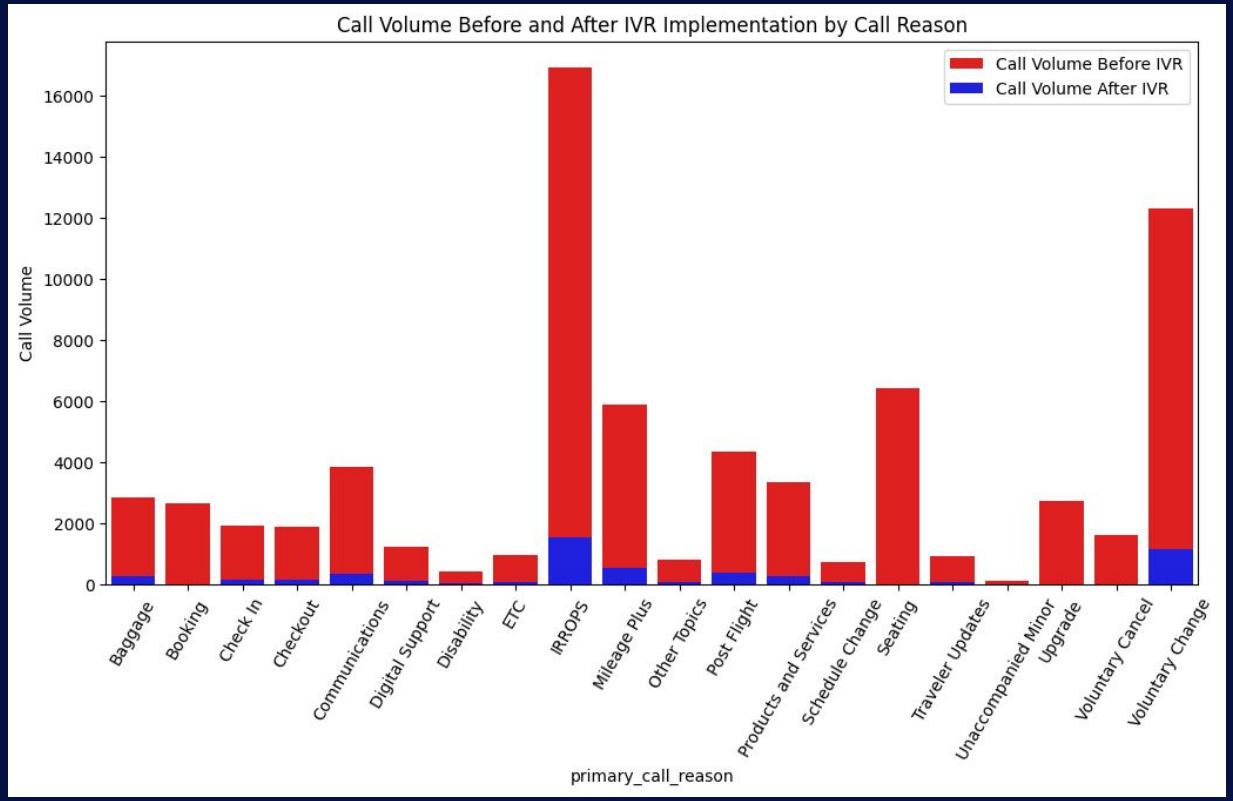



```
Topic 0:
['need', 'date', 'work', 'help', 'let', 'fee', 'customer', 'agent', 'change', 'flight']
Topic 1:
['booked', 'refund', 'just', 'travel', 'credit', 'let', 'change', 'agent', 'customer', 'flight']
Topic 2:
['delayed', 'like', 'sir', 'meeting', 'delay', 'let', 'tomorrow', 'customer', 'agent', 'flight']
Topic 3:
['help', 'weather', 'time', 'let', 'check', 'seat', 'wanted', 'customer', 'agent', 'flight']
Topic 4:
['let', 'united', 'delays', 'refund', 'voucher', 'experience', 'delay', 'flight', 'agent', 'customer']
Percentage of calls that could be solved via IVR: 93.24%
```

Recommendations for Optimizing Call Center Performance

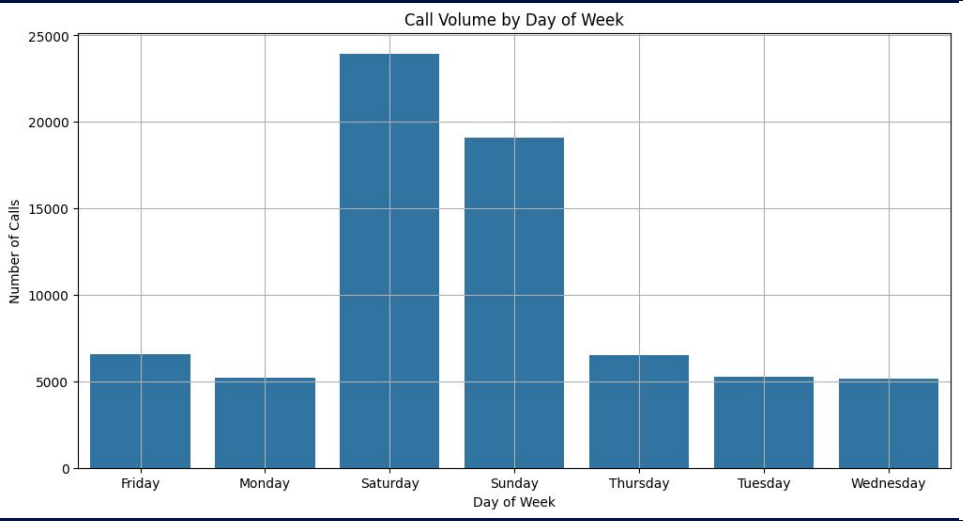
1. **Enhance IVR System** by providing self service options for Booking and Check-in, Flight Status and Baggage Tracking

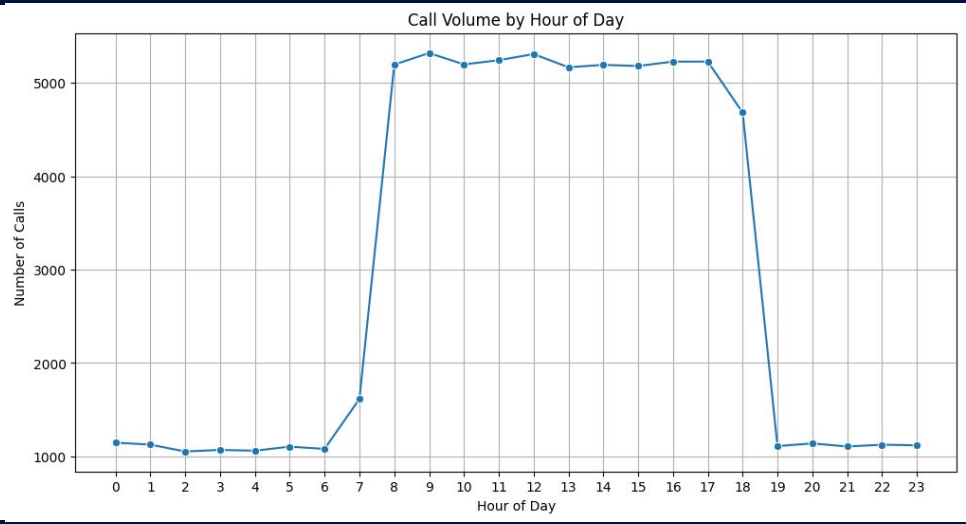
2. **Proactive Customer Communications** for common issues (e.g., flight delays, cancellations) via text or email to reduce call volume.



3. **Peak Hour Optimization and Agent Allocation:**

- Ensure more agents are available during peak times, such as early morning and weekdays.
- **Flexible Scheduling**: Use part-time or on-call agents to accommodate surges in call volume.
- More agents during peak hours ensures quicker response times and faster resolution.





Predicting Missing Call Reasons

- **Dataset Size**: 71,810 records
- **Missing Data**: 5,157 records with missing primary_call_reason, which we predicted using machine learning.
- **Clustering of Call Transcripts**: Grouped similar call types to identify common patterns and themes using NLP.
- **Predicting Missing Call Reasons**: Used TF-IDF for text vectorization and Random Forest for classification to predict missing primary_call_reason.
- Predicted call reasons are stored in Test.csv