# 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](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 provide **Self Solvable Options** and **Peak Hour Optimization**.

4. Predicted 5,157 Missing Call Reasons using 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%
```



Average Handle Time (AHT) for 5 Most and Least Frequent Call Reasons



Average Speed to Answer (AST) for 5 Most and Least Frequent Call Reasons
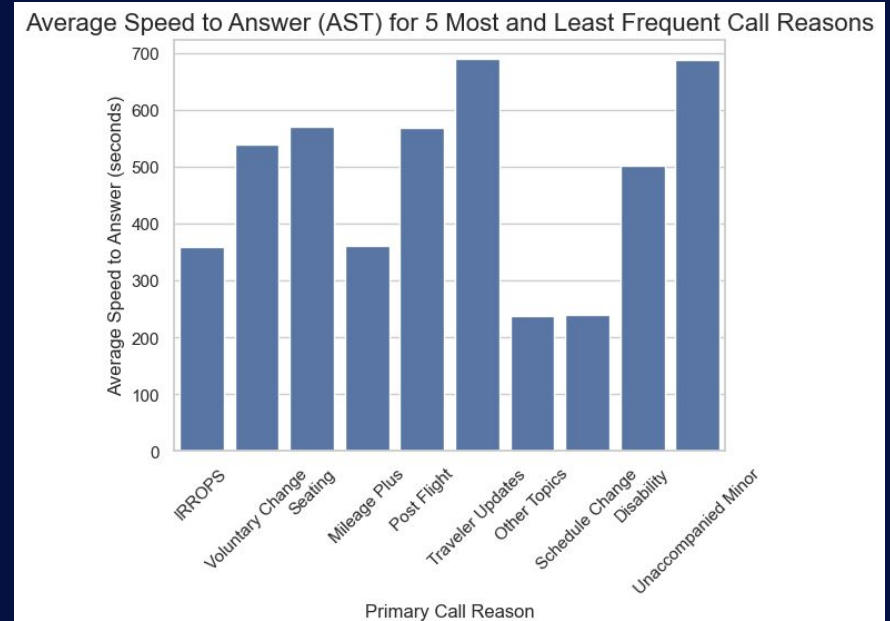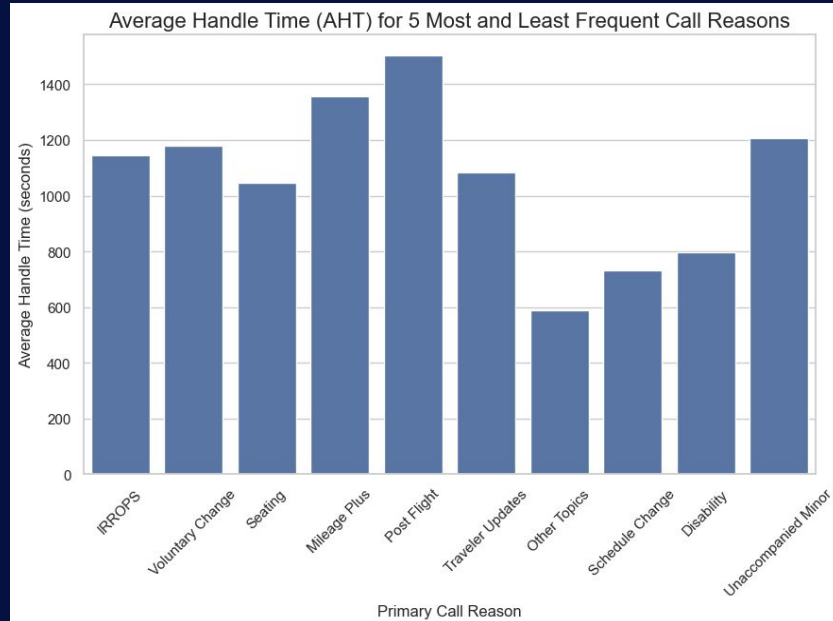
**2. Negative Average Sentiment**: Calls with High Negative Average Sentiment have longer handling times.

As shown in Relation Between Average Sentiment and Average Handling Time (AHT) Plot.

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



Correlation Heatmap of AHT, Tone, and Sentiment



Mean average_sentiment by Call Reason



Relationship Between Average Sentiment and Average Handle Time (AHT)
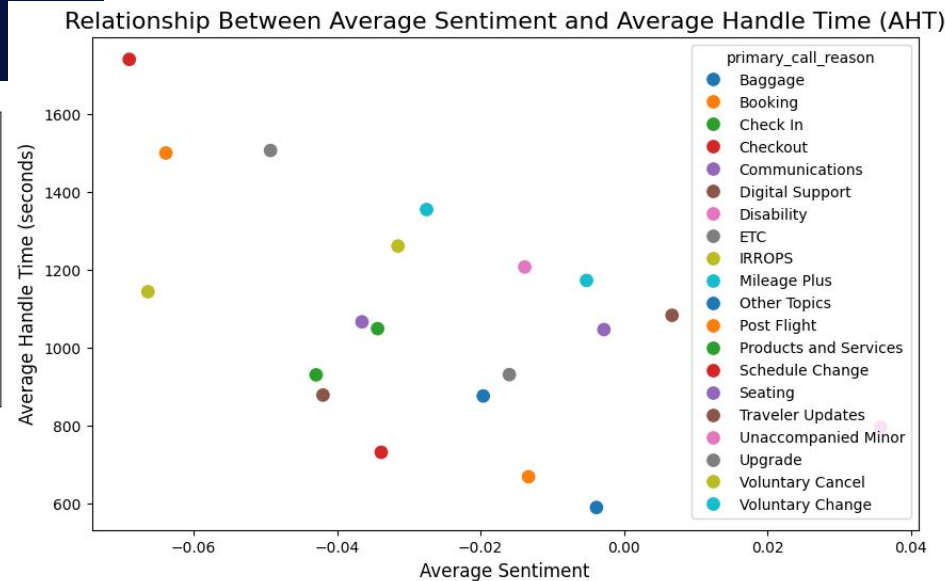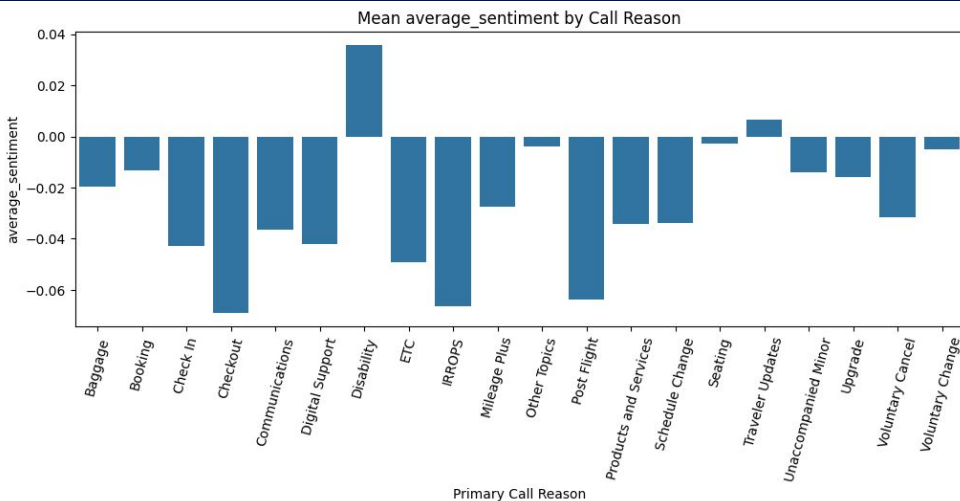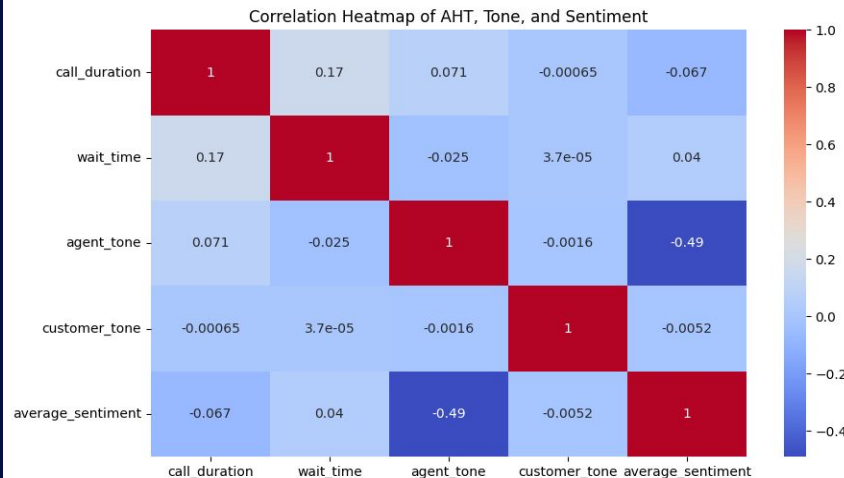
# Key Findings for self solvable issues

- **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.**



Word Cloud of Frequent Terms in Call Transcripts
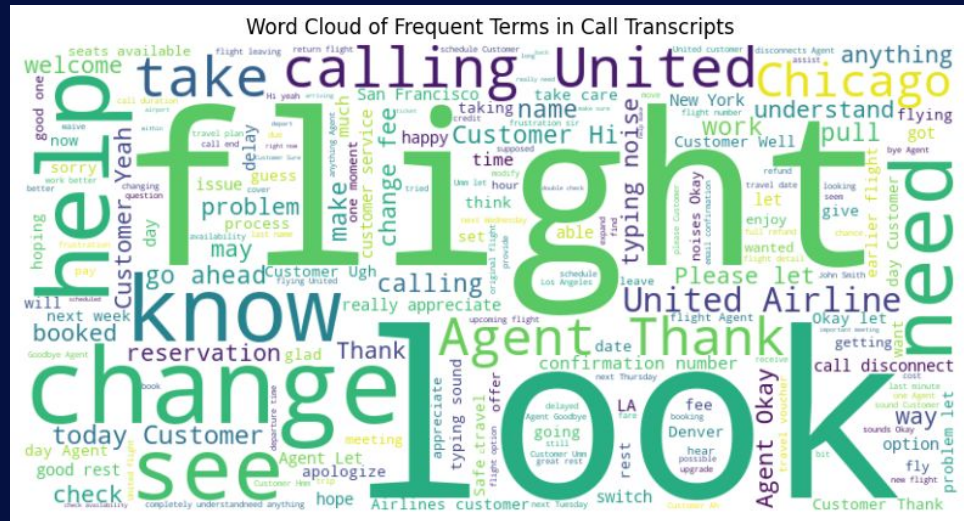
```
Topic 0:
['need', 'date', 'work', 'help', 'let', 'fee', 'customer', 'agent', 'change', 'flight']
Topic 1:
['booked', 'refund', 'just', 'travel', 'credit', 'let', 'change', 'agent', 'customer', 'flight']
Topic 2:
['delayed', 'like', 'sir', 'meeting', 'delay', 'let', 'tomorrow', 'customer', 'agent', '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. **Proactively Communicate Customer** for common issues (e.g., flight delays, cancellations) via text or email to reduce call volume.



Call Volume Before and After IVR Implementation by Call Reason
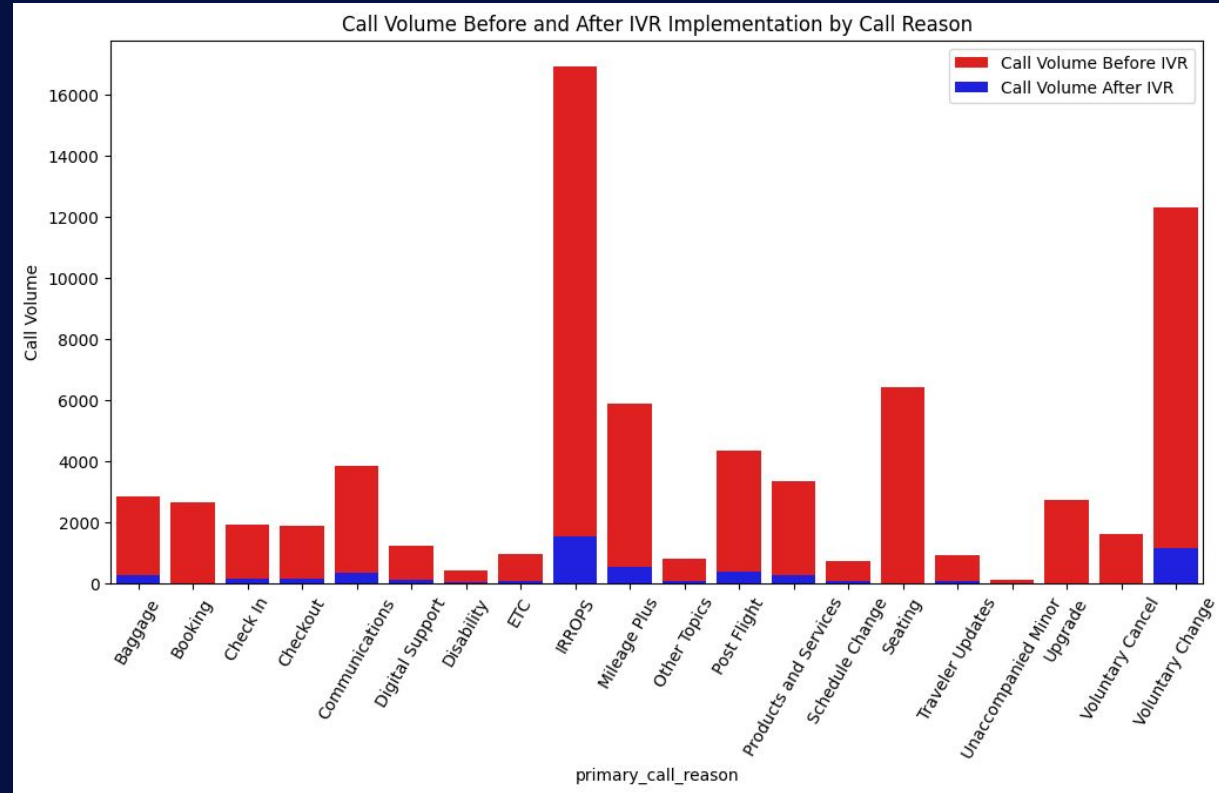
## 3. **Peak Hour Optimization and Agent Allocation:**

- 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.
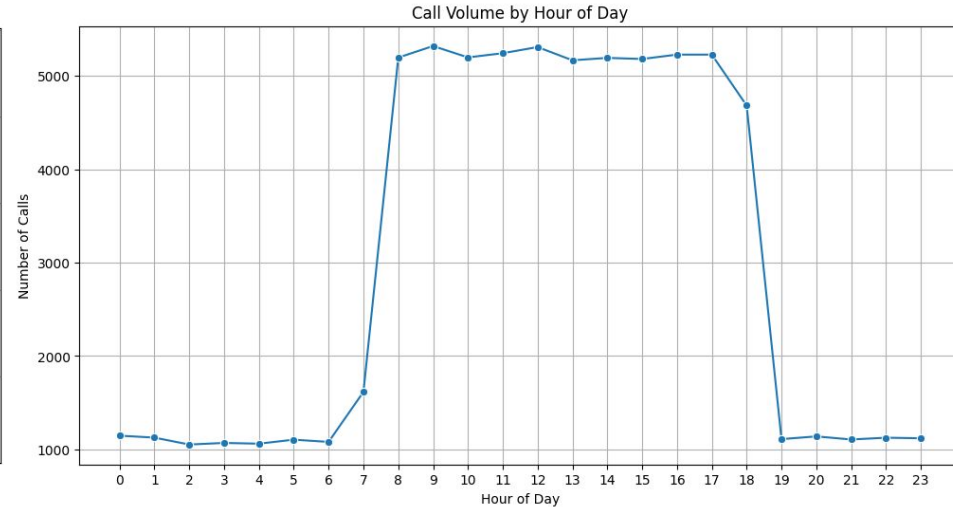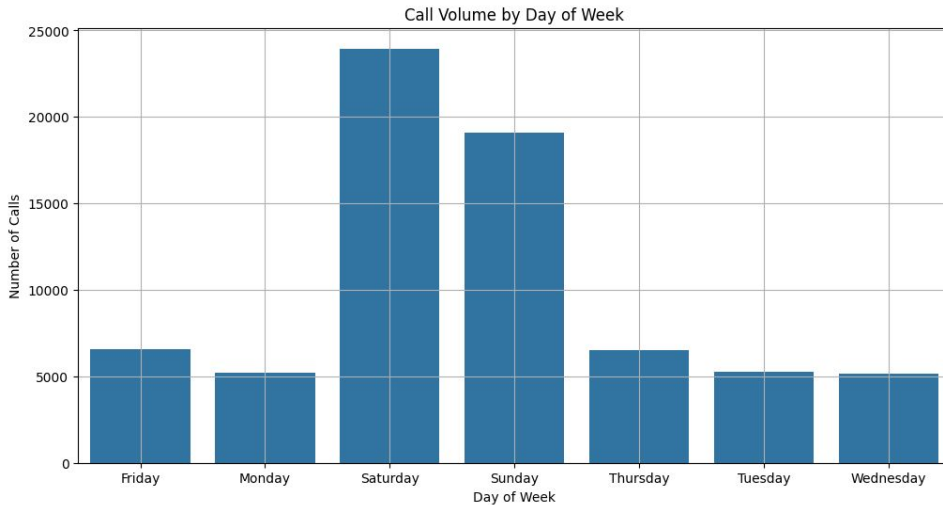
# Predicting Missing Call Reasons

- **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



Correlation Between Call Reasons and Identified Clusters (Recurring Issues)

| primary_call_reason | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| | 2.5e+03 | 5.5e+02 | 7.8e+02 | 2.9e+02 | 1.1e+03 |
| Baggage | 1.5e+03 | 2.7e+02 | 3.8e+02 | 1.3e+02 | 6e+02 |
| Booking | 1.3e+03 | 2.7e+02 | 3.7e+02 | 1.2e+02 | 5.4e+02 |
| Check In | 9.2e+02 | 2.1e+02 | 2.9e+02 | 1.1e+02 | 3.7e+02 |
| Checkout | 9.1e+02 | 2e+02 | 2.7e+02 | 1.1e+02 | 4e+02 |
| Communications | 1.9e+03 | 3.9e+02 | 5.6e+02 | 2.2e+02 | 7.5e+02 |
| Digital Support | 5.6e+02 | 1.4e+02 | 2.2e+02 | 58 | 2.5e+02 |
| Disability | 1.7e+02 | 53 | 69 | 20 | 90 |
| ETC | 4.3e+02 | 1e+02 | 1.5e+02 | 63 | 2e+02 |
| IRROPS | 6.5e+03 | 1.6e+03 | 1.7e+03 | 8.4e+02 | 2.7e+03 |
| Mileage Plus | 2.9e+03 | 5.4e+02 | 9.2e+02 | 3.5e+02 | 1.2e+03 |
| Other Topics | 3.8e+02 | 66 | 1.5e+02 | 54 | 1.7e+02 |
| Post Flight | 2e+03 | 4.4e+02 | 6.4e+02 | 2.5e+02 | 1e+03 |
| Products and Services | 1.5e+03 | 3.4e+02 | 5e+02 | 1.8e+02 | 7.9e+02 |
| Schedule Change | 3.4e+02 | 89 | 1.2e+02 | 37 | 1.5e+02 |
| Seating | 3e+03 | 6.7e+02 | 1e+03 | 3.7e+02 | 1.3e+03 |
| Traveler Updates | 4.4e+02 | 95 | 1.6e+02 | 51 | 1.9e+02 |
| Unaccompanied Minor | 43 | 15 | 18 | 8 | 20 |
| Upgrade | 1.3e+03 | 3e+02 | 4.3e+02 | 1.9e+02 | 5.5e+02 |
| Voluntary Cancel | 7.6e+02 | 1.4e+02 | 2.5e+02 | 1.1e+02 | 3.4e+02 |
| Voluntary Change | 5.2e+03 | 1.1e+03 | 1.8e+03 | 5.8e+02 | 2.2e+03 |

cluster