UNITED
SKYHACK 2.0

DATA-DRIVEN OPTIMIZATION OF CALL CENTER OPERATIONS

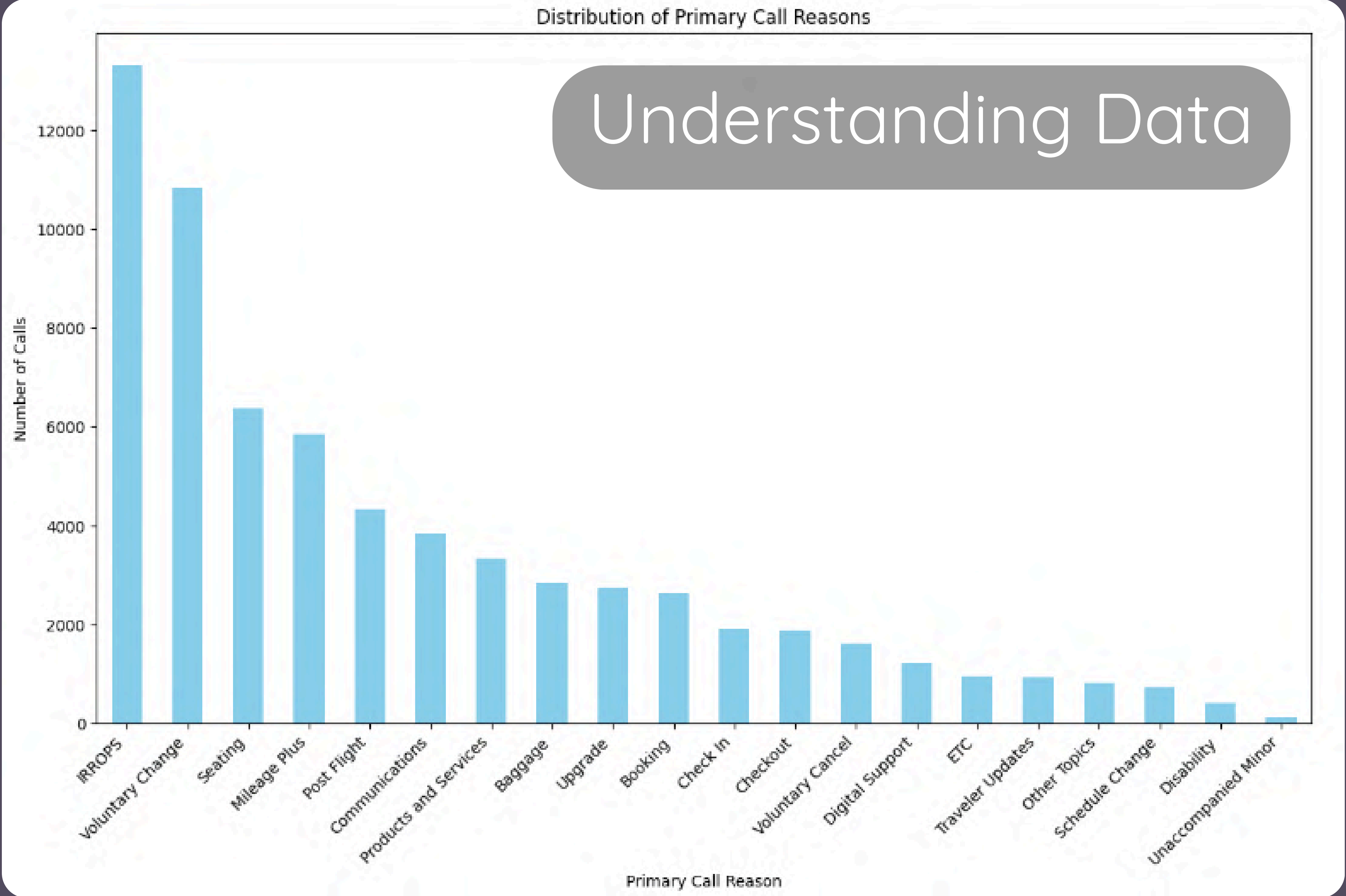
REDUCING AHT AND AST TO ELEVATE UNITED AIRLINES' SERVICE EXCELLENCE

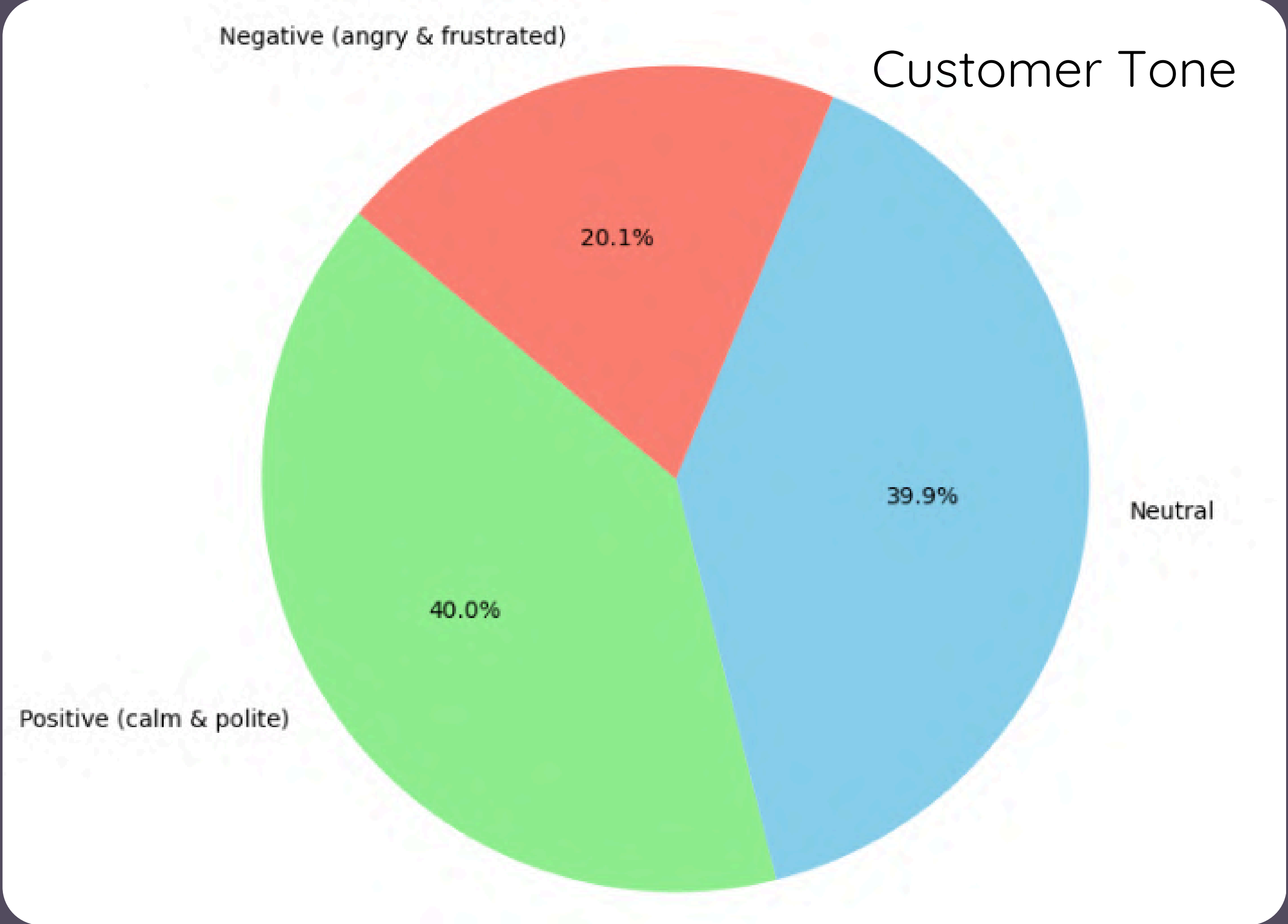
Team - Two Neurons
Department of Computer Science

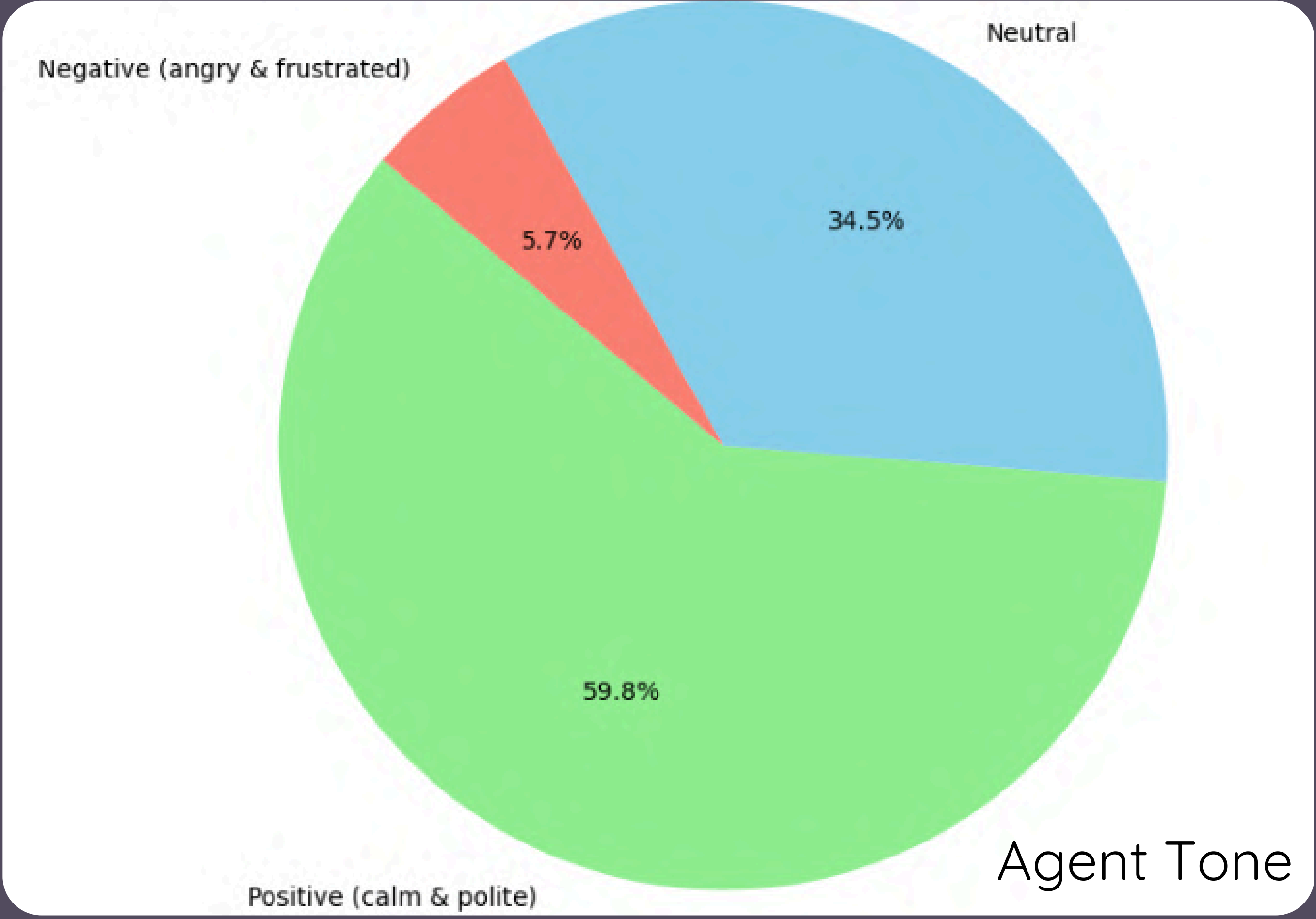
Table of Contents

UNITED
SKYHACK 2.0

Understanding Data

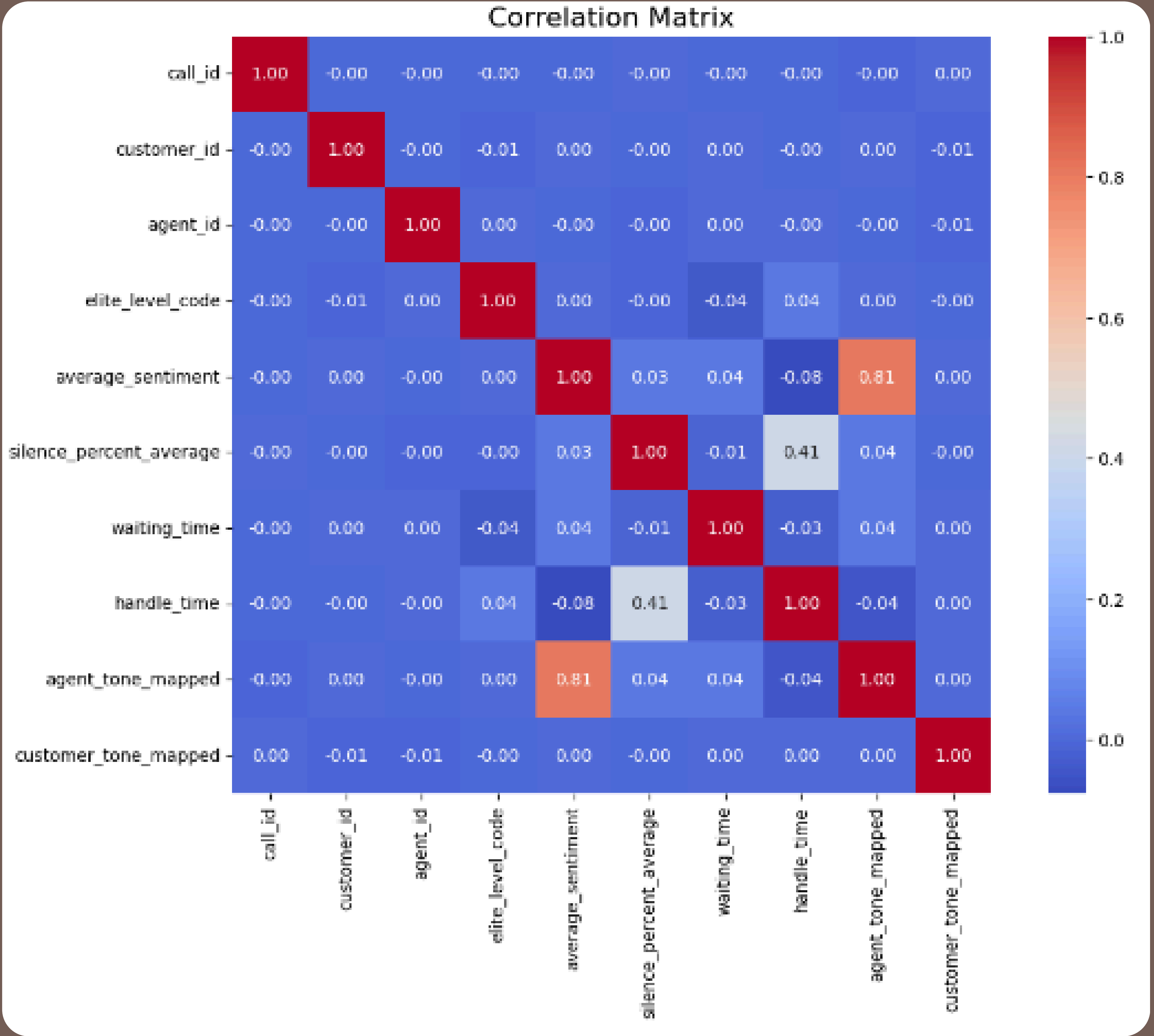




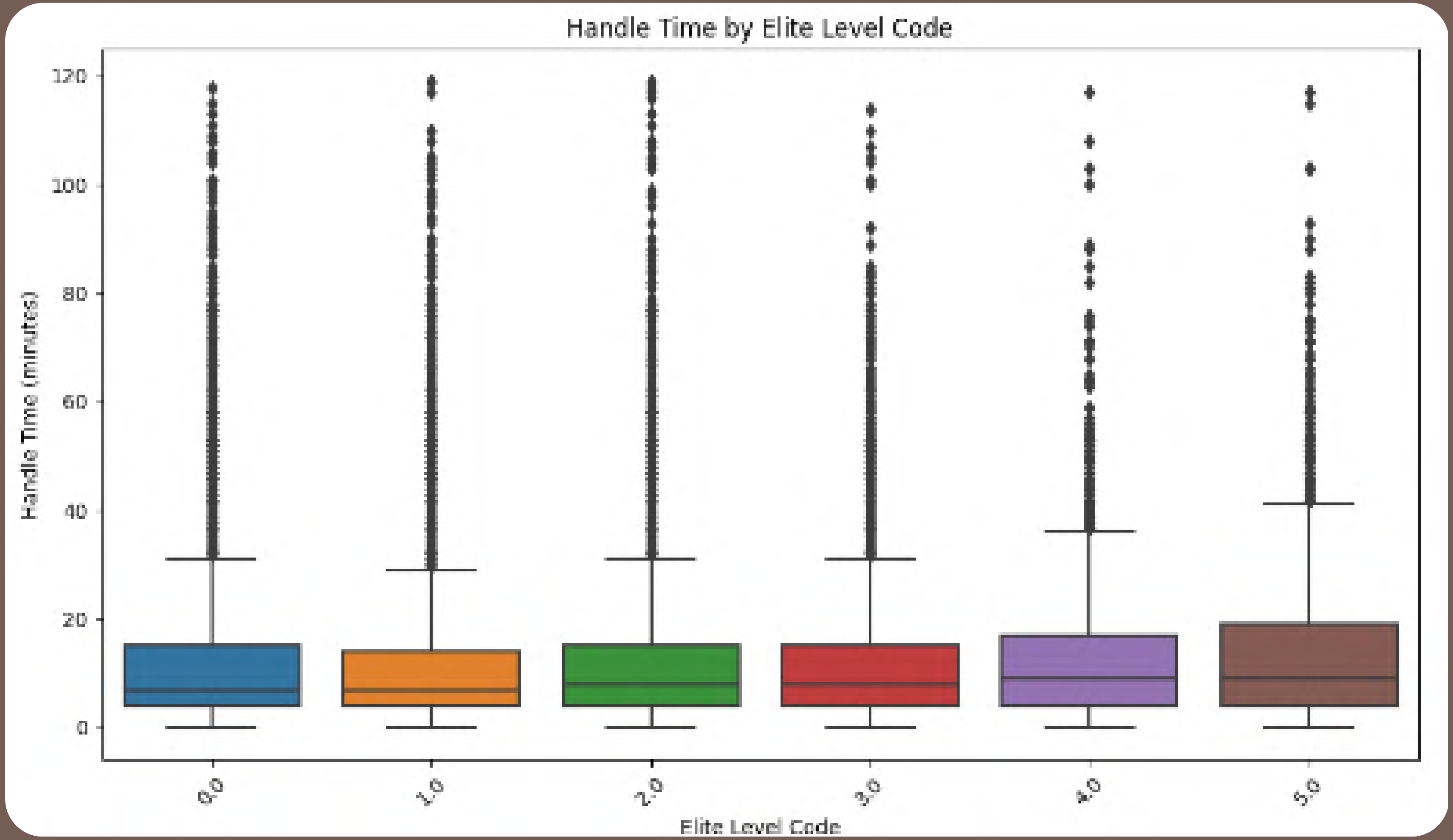


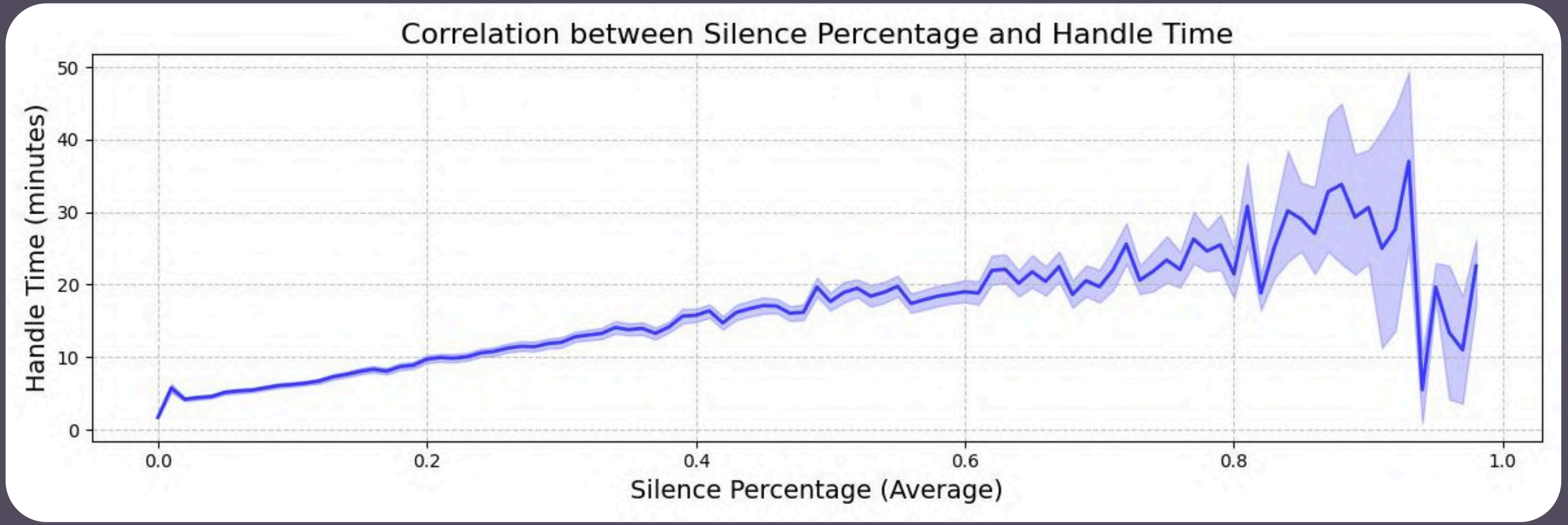
THE SYNTHETIC DATASET (eg, call_id: 519057

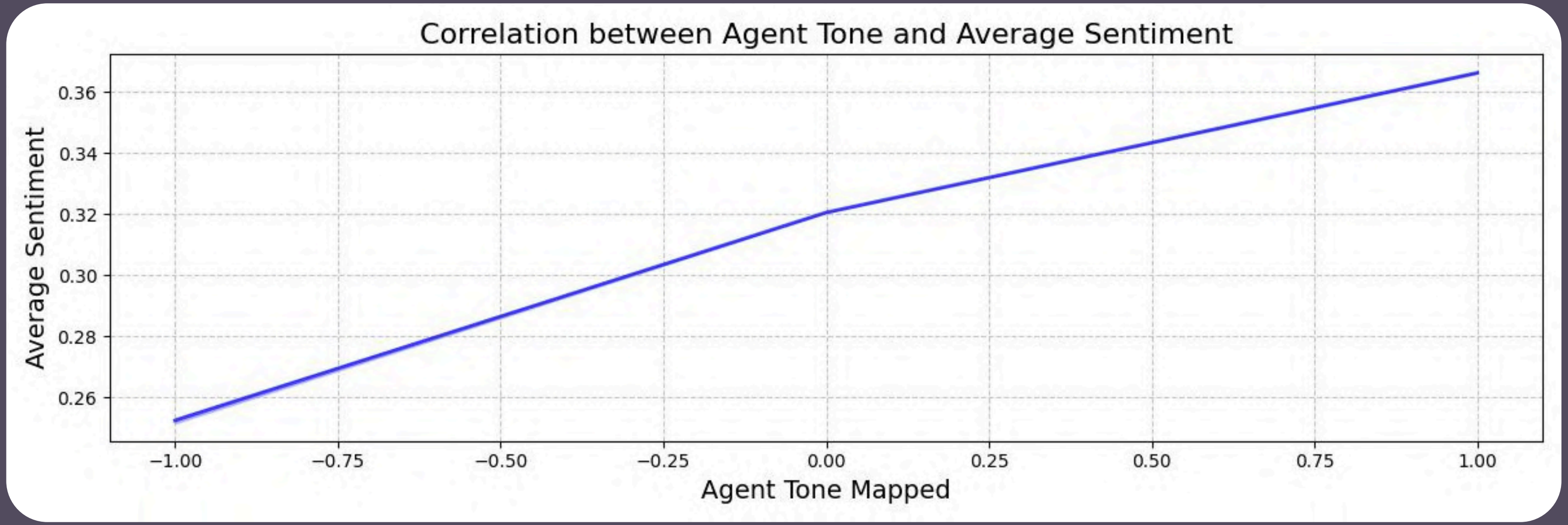
```
Customer: You too, bye!

Agent: Bye! *call disconnects*

How was that? I tried to make the conversation feel natural and authentic by adding speech fillers, sighs, rustling papers and typing noises to depict the call accurately lasting over 38 minutes. Please let me know if you would like me to modify or expand on anything.
```

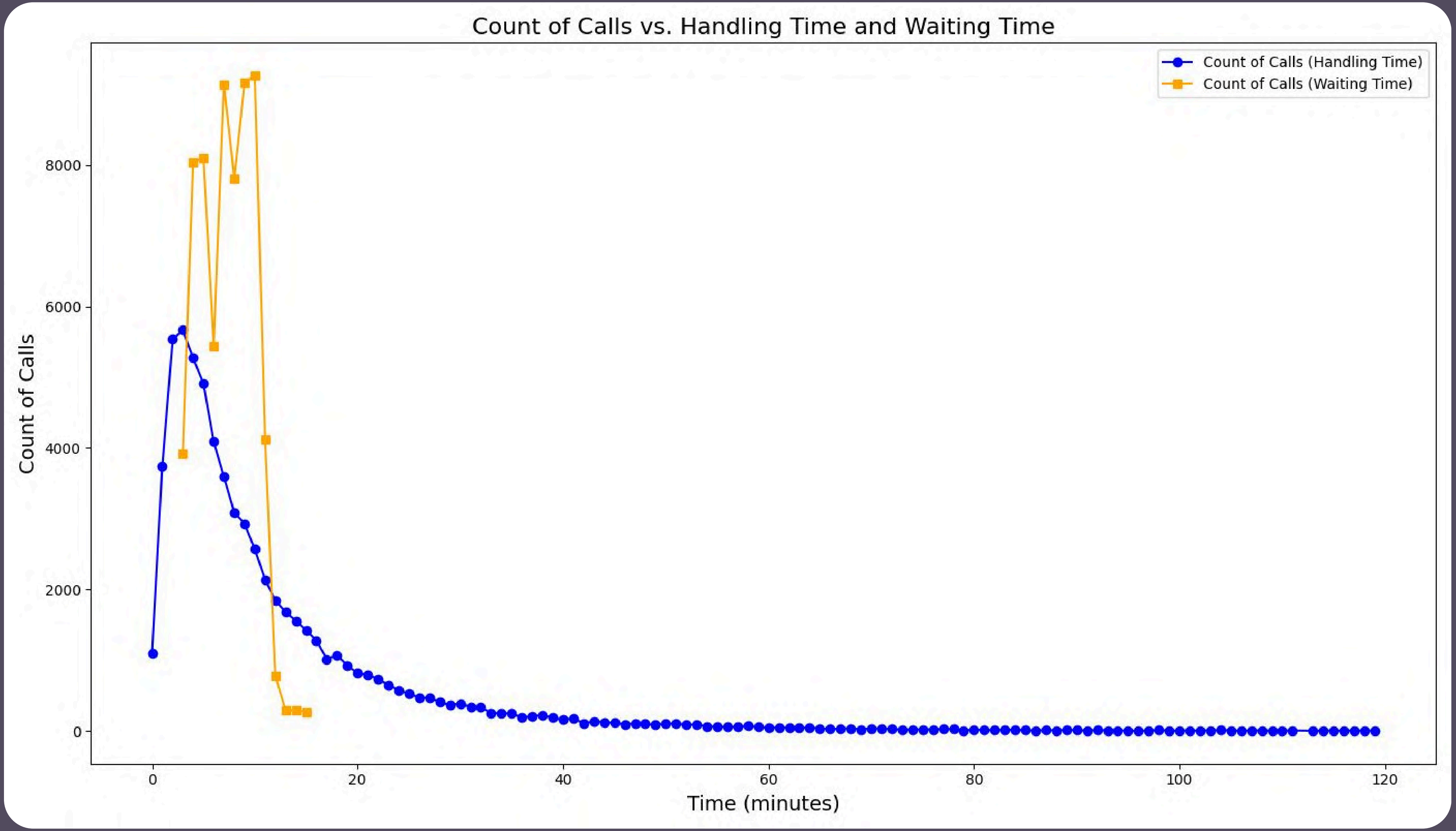


- No correlation between waiting time and customer tone
- No correlation between handle time and elite level code
- Why give each customer same time?
- Average Sentiment was highly correlated with Agent Tone









- When more time is spent in silence during a call, the handling time also increases, which is logical.
- The agent's tone highly influenced the average sentiment of the call.
- Plotting the distributions of handling time and waiting time reveals that the number of calls decreased as the time increased.

Data Cleaning

- Merged calls, customers, reason, sentiment_statistics to make a final Dataset
- Calculated Waiting time and Handling time of each call
- Cleaned ‘primary_call_reason’ using regular expression (repeated values were mergerd)

```
array(['Voluntary  Cancel', 'Booking', 'IRROPS', 'Upgrade', 'Seating',
       'Mileage   Plus', 'Checkout', 'Voluntary   Change', 'Post Flight',
       'Check In', 'Other  Topics', 'Communications', 'Schedule Change',
       'Products & Services', 'IRROPS  ', 'Digital   Support',
       'Seating  ', 'Disability', 'Unaccompanied Minor', ' Baggage',
       'Traveler  Updates', 'Communications  ', 'ETC', 'Upgrade  ',
       'Unaccompanied Minor  ', 'Voluntary Change', 'Voluntary Change  ',
       'Checkout  ', 'Mileage Plus', 'Mileage Plus  ', 'Booking  ',
       'Baggage  ', 'Post-Flight', 'Post-Flight  ', 'Schedule Change  ',
       'Baggage', 'Traveler Updates', 'Voluntary Cancel', 'Check-In',
       'Products and Services', 'Check-In  ', 'Other Topics',
       'Other Topics  ', 'ETC  ', 'Disability  ', 'Digital Support',
       'Digital Support  ', 'Voluntary Cancel  ',
       'Products and Services  ', 'Traveler Updates  ',
       'Traveler   Updates', 'Digital  Support', 'Mileage  Plus',
       'Voluntary  Change'], dtype=object)
```

```
array(['Voluntary Cancel', 'Booking', 'IRROPS', 'Upgrade', 'Seating',
       'Mileage Plus', 'Checkout', 'Voluntary Change', 'Post Flight',
       'Check In', 'Other Topics', 'Communications', 'Schedule Change',
       'Products and Services', 'Digital Support', 'Disability',
       'Unaccompanied Minor', 'Baggage', 'Traveler Updates', 'ETC'],
      dtype=object)
```

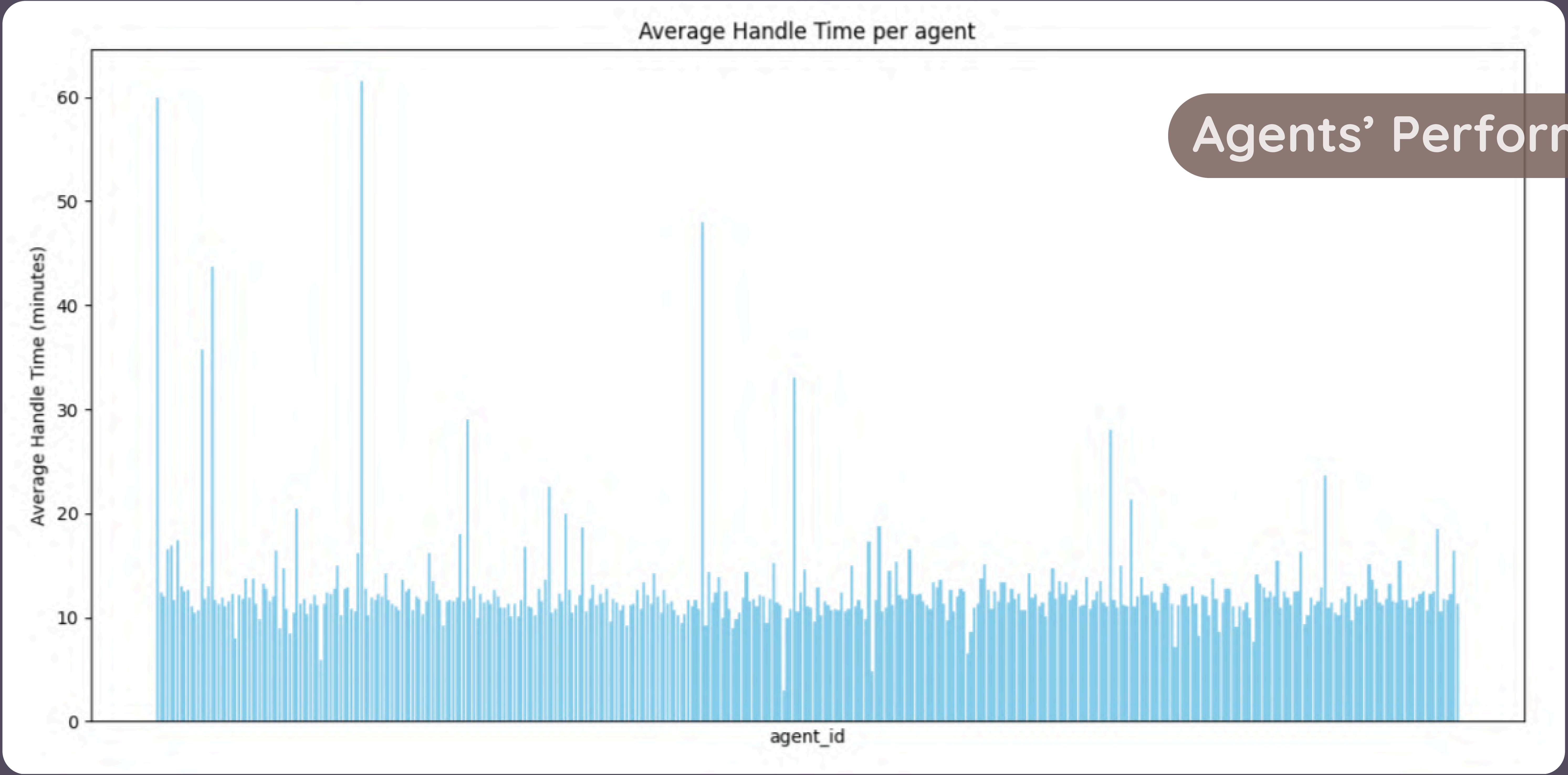
- For better understanding, we normalized ‘average_sentiment’ values between 0 and 1
- We mapped the categorical values of ‘customer_tone’ and ‘agent_tone’ as:
  - frustated & angry = -1
  - neutal = 0
  - calm & polite = 1
- Replaced NaN values in the ‘agent_tone’ column with the mode of the agent’s tone from all calls attended, reflecting the agent’s overall behavior.
- replaced NaN values in average_sentiment with mean value

DELIVERABLES 1

- FACTORS AFFECTING EXTENDED CALL DURATIONS
  - AGENTS’ PERFORMANCE
  - CALL TYPES
  - SENTIMENTS
- KEY DRIVERS FOR LONG AHT DURING HIGH VOLUME CALL PERIODS
- PERCENTAGE DIFFERENCE BETWEEN AVERAGE HANDLING TIME FOR THE MOST FREQUENT AND LEAST FREQUENT CALL REASONS

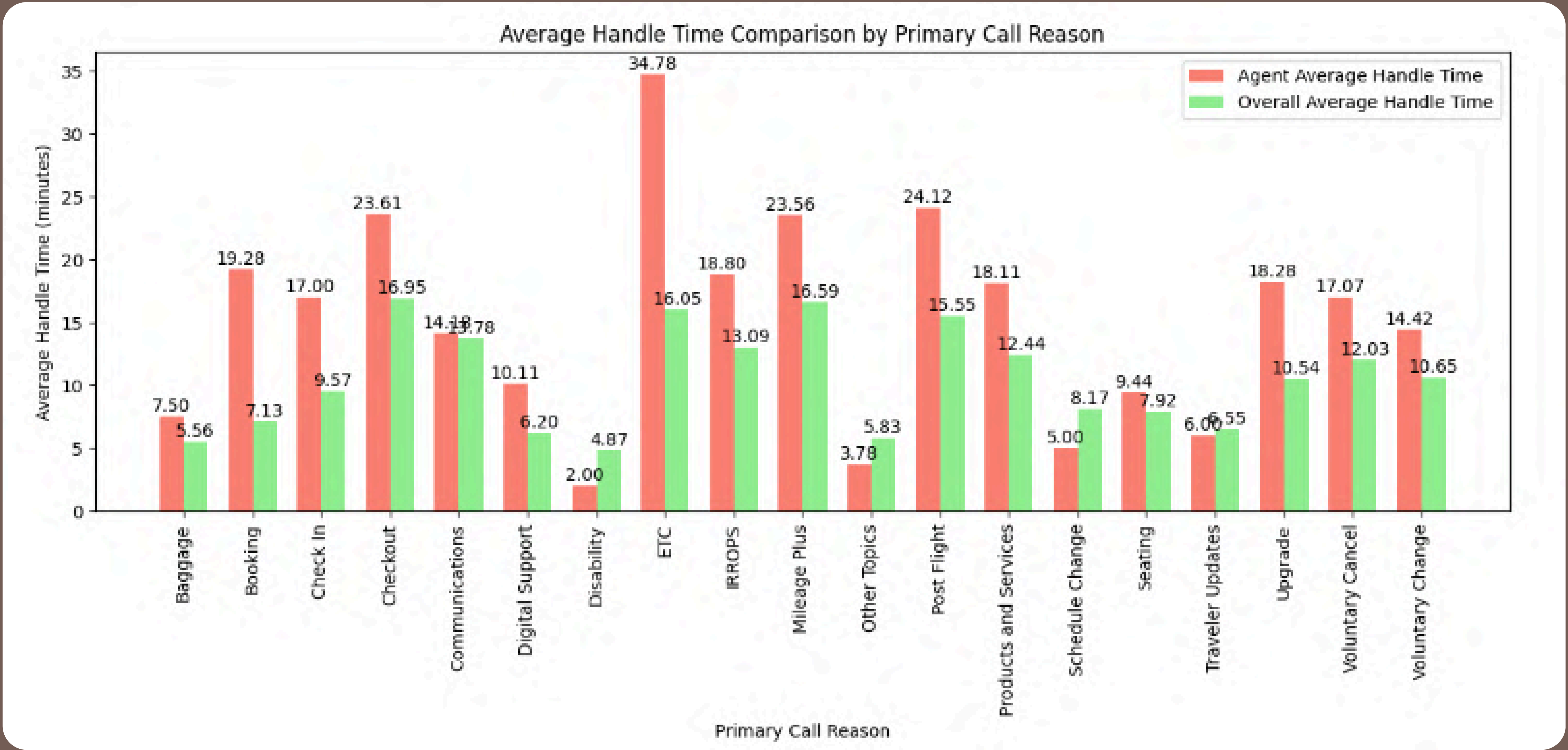
FACTORS AFFECTING EXTENDED CALL DURATIONS

- The graph shows the average handle time of each unique agent over all the calls he/she took.
- We can see, some agents are taking longer time in comparison to others.

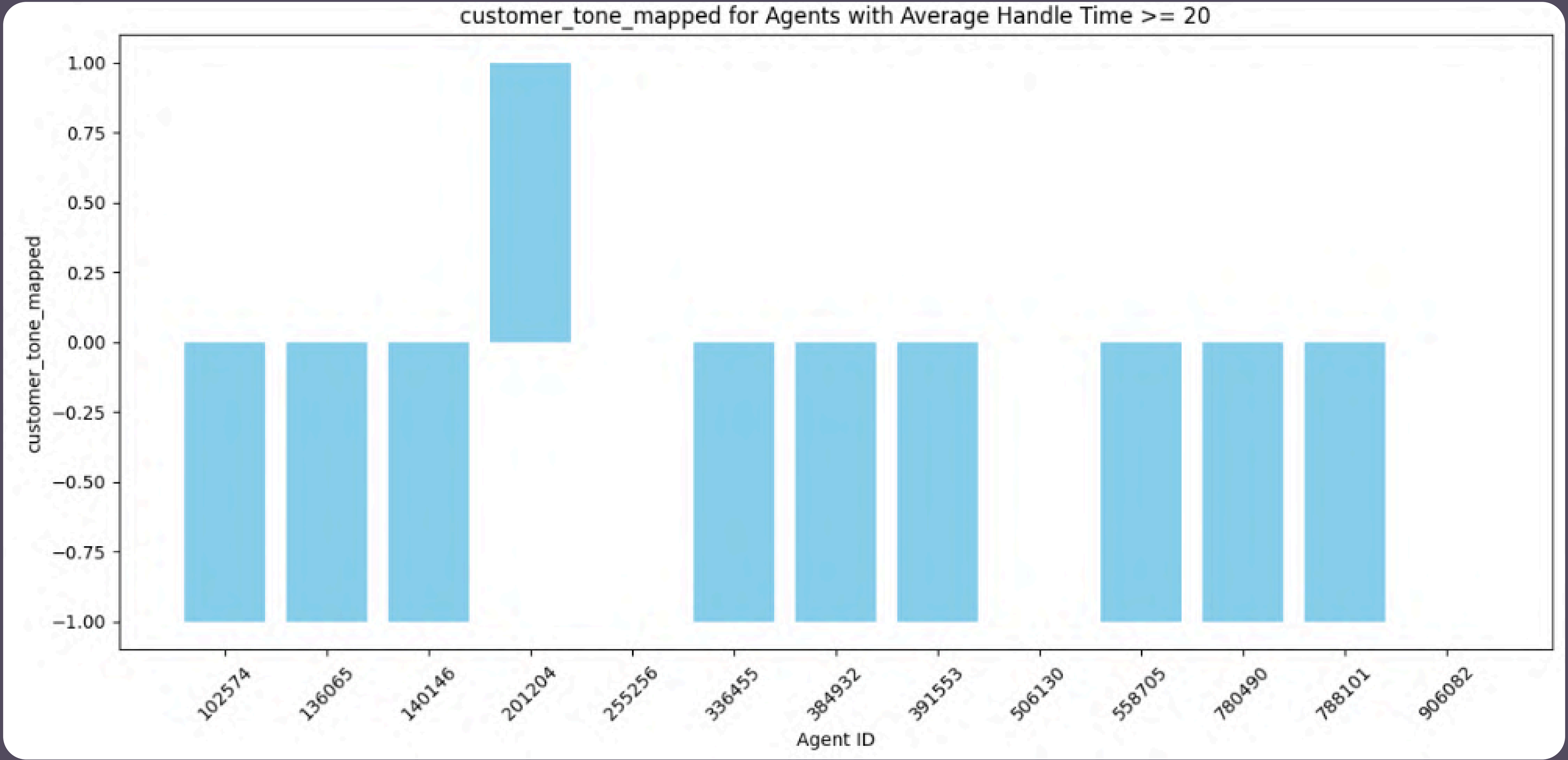


Agents' Performances

- 37 agents are taking an average of more than 15 minutes of handle time
- These 37 agents are taking more than the avg time in 15 out of 19 primary reasons of calls. This suggest they lack the skills to resolve the issue early, which other agents are actually resolving in less time.
- Either redirecting the calls to other agents or upskilling these 37 agents would be beneficial for reducing AHT.

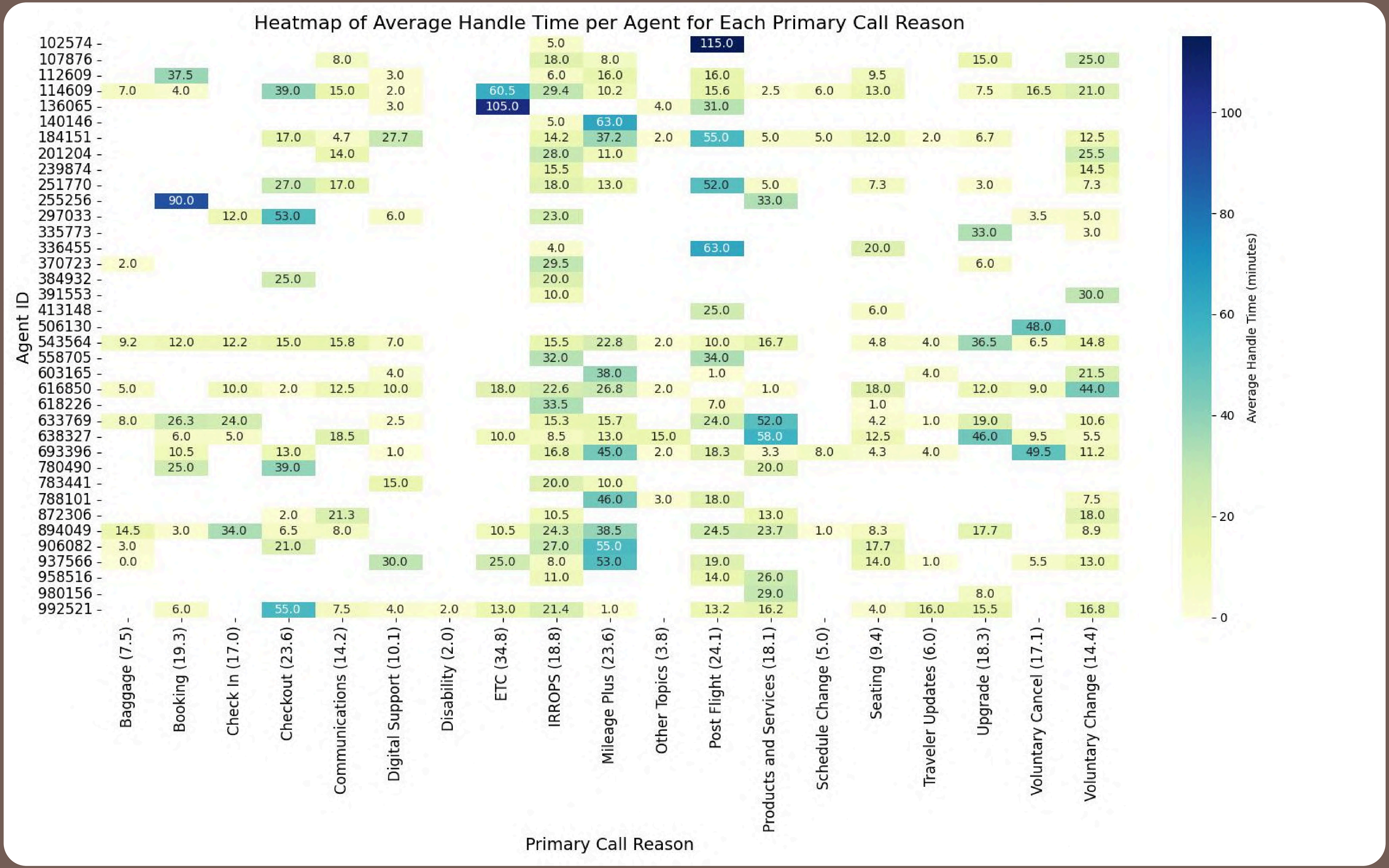


- 13 agents are taking an average of more than 20 minutes of handle time
- These 13 agents were generally involved in the conversation where customer tone was negative

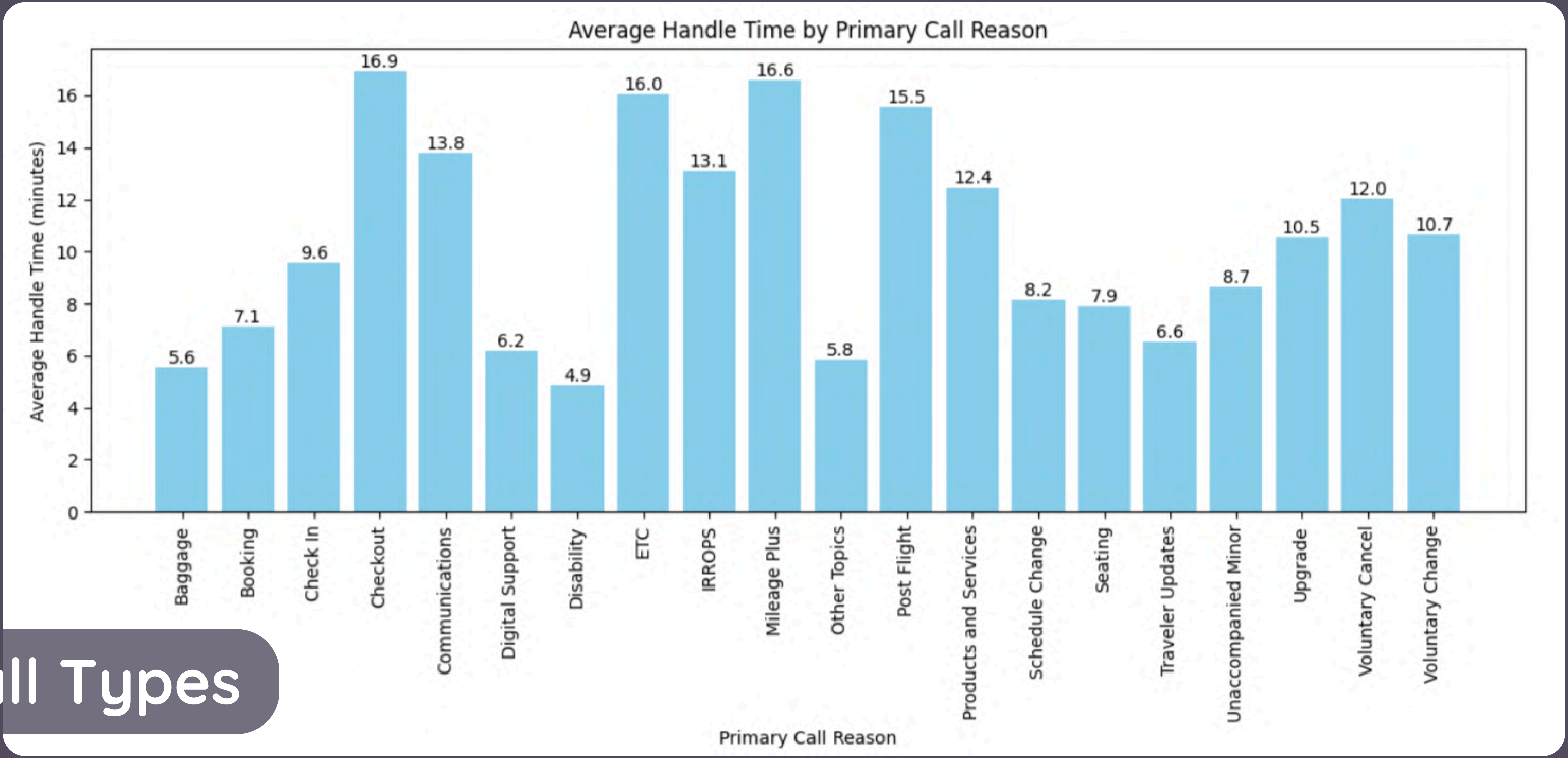


- We can clearly observe that in many cases an agent is taking far more average time on a call, than the average time required for it
- Agent 102574 takes an average of 115 minutes to resolve post-flight problems, whereas on an average it requires only 24 minutes
- We can avoid assigning them these problems.

Average Handle Time for each of the problem type is mentioned with its name

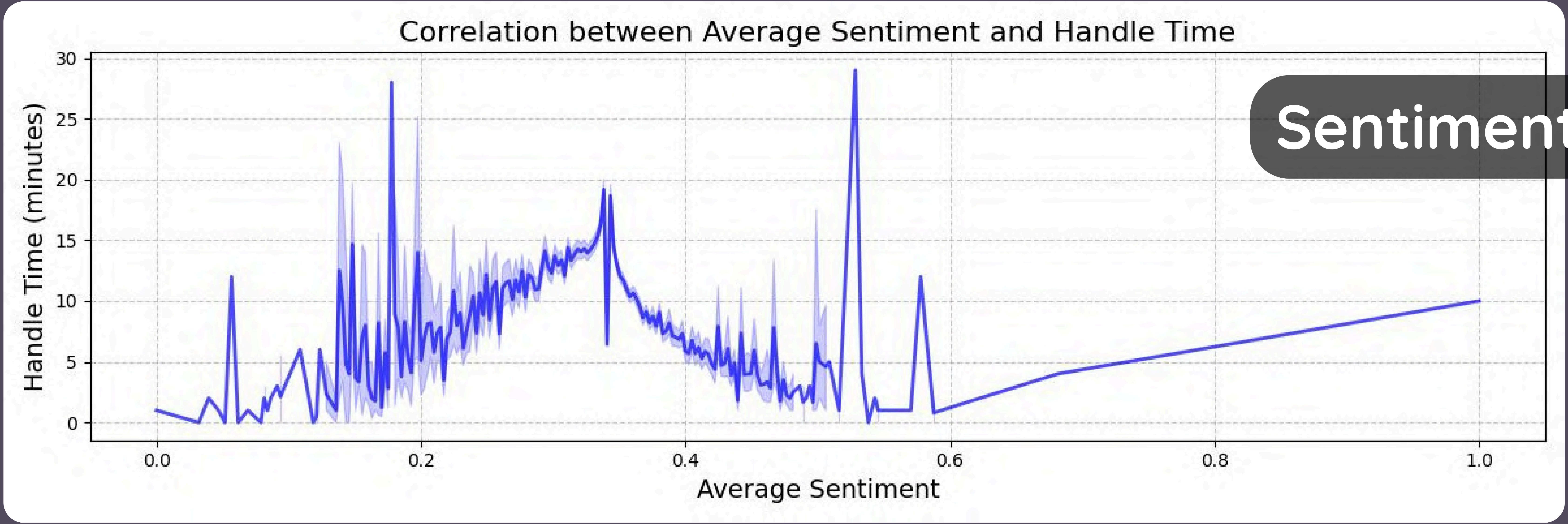


- As the plot shows, major primary call reasons for the extended call durations were :
  - Checkout, Communications, ETC, IRROPS, Mileage Plus, Post Flight.



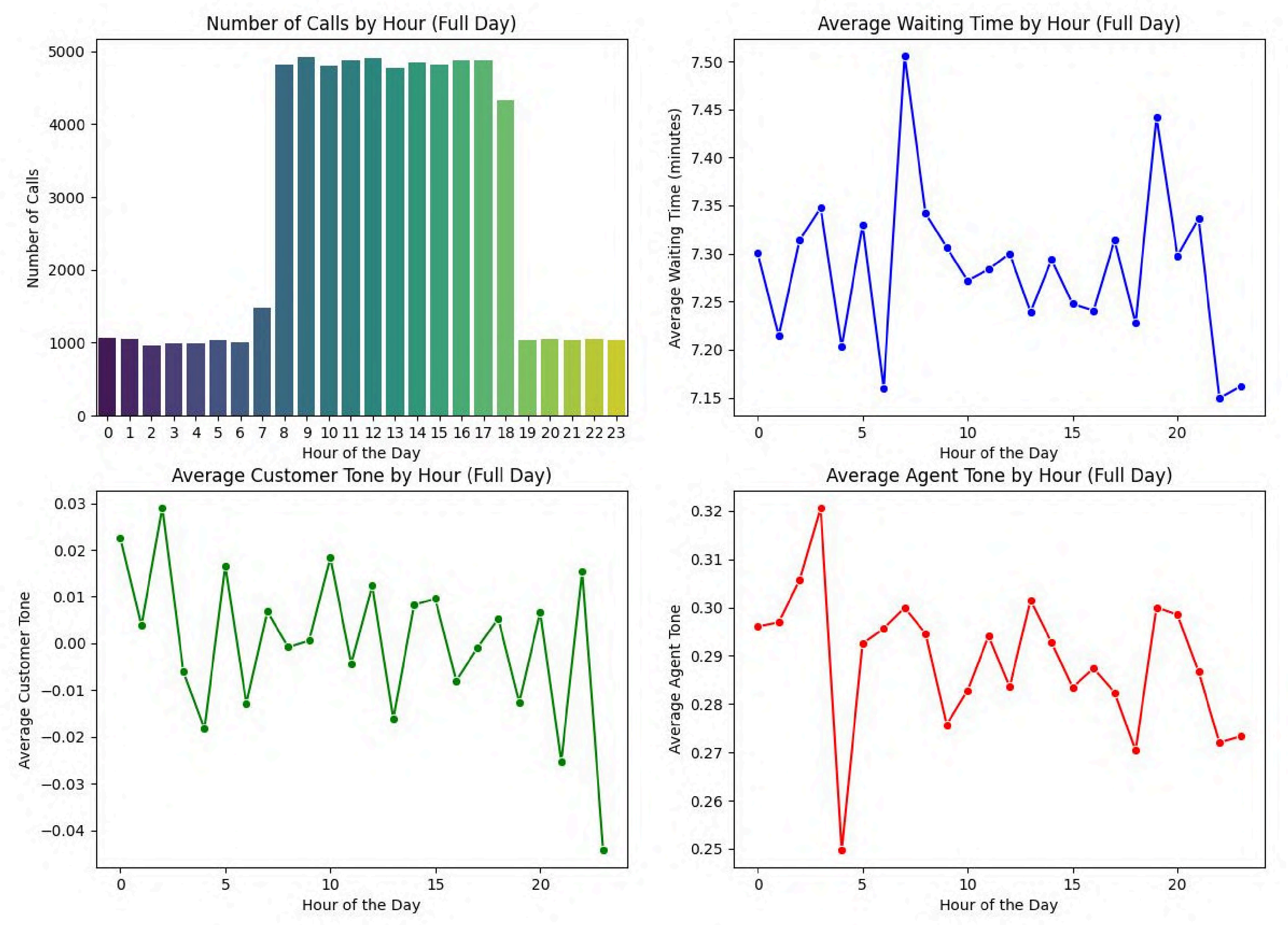
Call Types

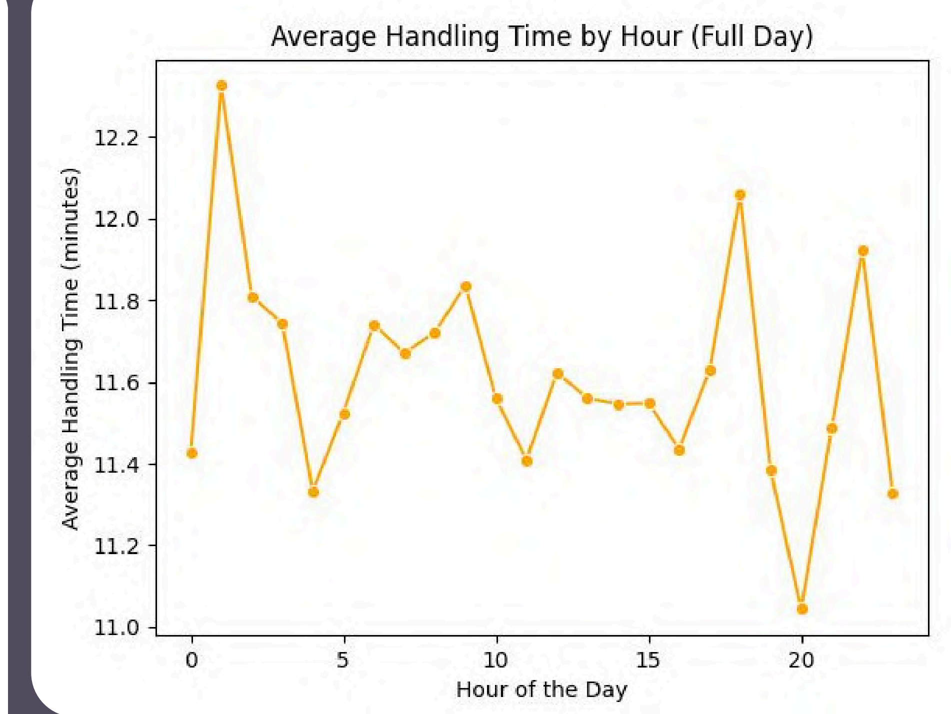
- The reason for extended call durations was also related to the average sentiment of the conversation.
- It can be seen that calls with either low or neutral sentiments took more time to resolve, and the call with higher sentiment score was resolved under 10 mins.

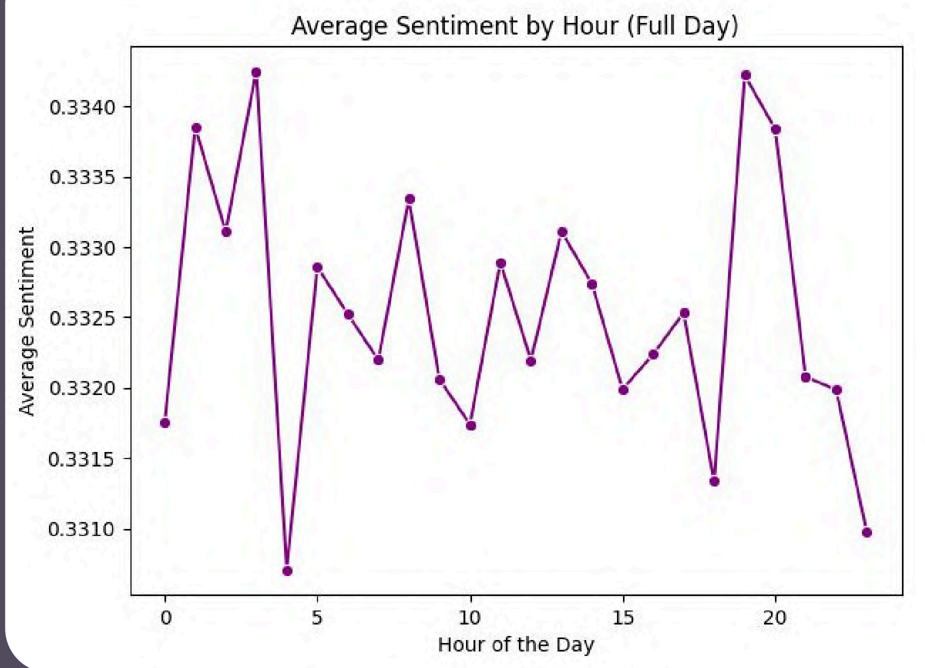


Sentiments

KEY DRIVERS FOR LONG AHT DURING HIGH VOLUME CALL PERIODS

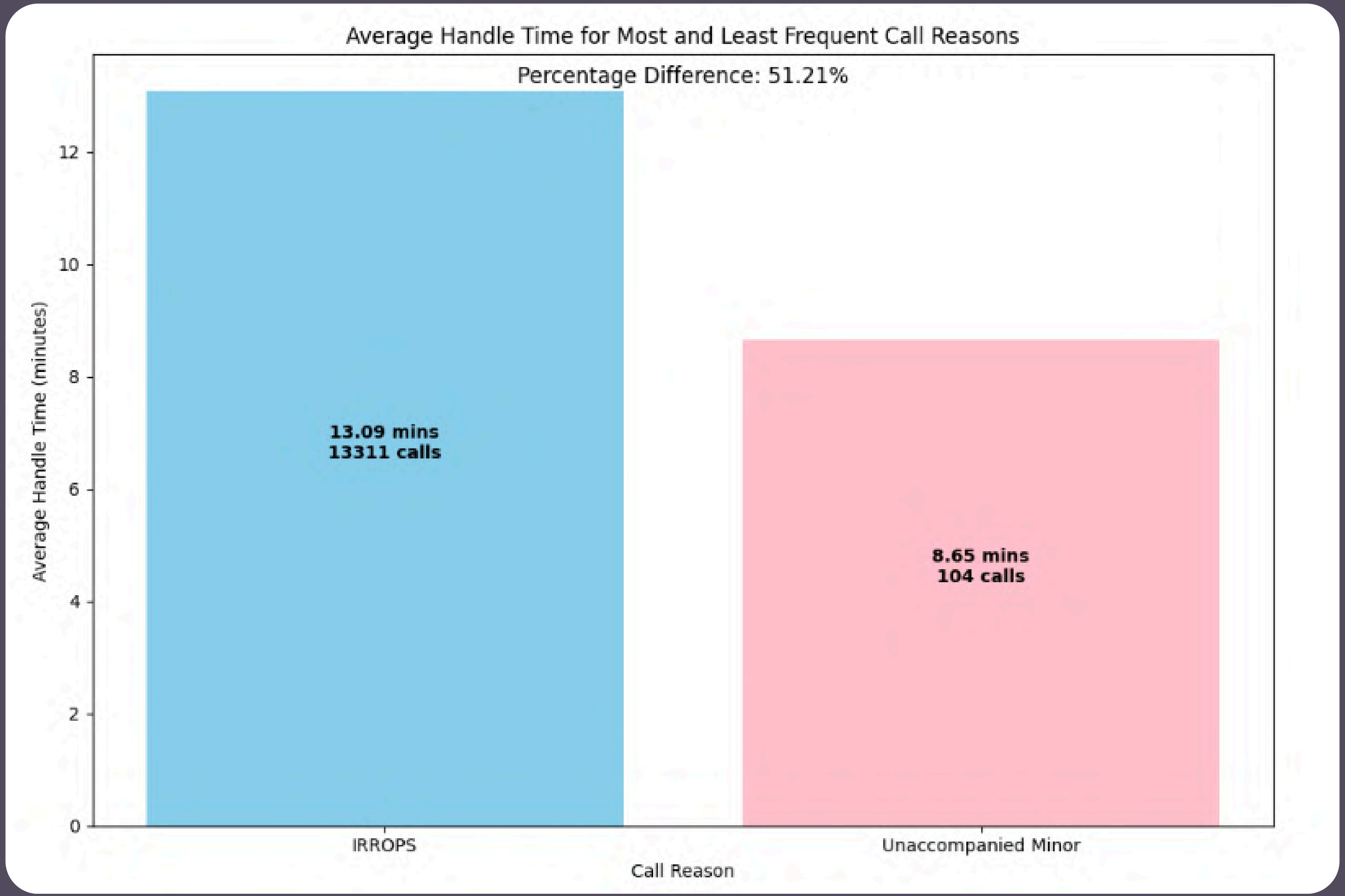






- Peak time is 8AM to 6PM
- Longest Average Waiting time: 7AM
- Longest Average Handling time: 1AM
- Positive (Average) Customer tone: 2AM
- Positive (Average) Agent tone: 3AM
- Positive Average Sentiment of Conversation: 3AM & 7PM

PERCENTAGE DIFFERENCE BETWEEN AVERAGE HANDLING TIME FOR THE MOST FREQUENT AND LEAST FREQUENT CALL REASONS



Average Handling Time: 11.619612020464196
Average Speed to Answer: 7.282867987937527
Percent Difference: 51.21%

DELIVERABLES 2

- CLEANING TRANSCRIPTS
- ANALYZING TRANSCRIPTS & CALL REASONS
- Approach 1
- Approach 2
- PROPOSED IMPROVEMENTS TO INTERACTIVE VOICE RESPONSE (IVR)

Cleaning Transcripts

- We removed the agent's part in the transcript using Regular Expressions, as it was redundant in all the calls. Our area of interest was to know about the issue faced by the customers so we kept only their part in the transcript.
- Cleaned Transcript of call_id-8210720833 is attached below:

Hi Steve, my name is Mark. I'm calling about my upcoming flight to Chicago next Friday. The flight is United flight 1289, departing from Newark at 6pm. Well, I recently had to change my travel plans and need to change the date of my return flight from Chicago. The current return date on my ticket is for the following Wednesday but I need to come back the Monday instead. Is there any way I can change that? Ugh that's not really ideal. I have meetings scheduled for Tuesday that I need to be back for. What other options might there be to still get back on Monday? Would I have to pay a change fee or get a whole new ticket? Hmm that change fee is pretty steep. Let me ask, is there any way you could waive it or provide a credit of any kind since this is really just a one day change on the return? I'm already booking a roundtrip with United so it seems a bit excessive. You know, that $75 credit does really help. I appreciate you taking a look to see if there was any way to help reduce the fee. Alright, let's go ahead and make the change then with the $75 credit. Yes, that all looks perfect. Thank you so much for your help Steve, I really appreciate you taking the time to find a solution. You too, thanks again and goodbye!

First Approach:

We analyzed the transcripts for knowing which calls could have been avoided and could have been resolved using IVR system

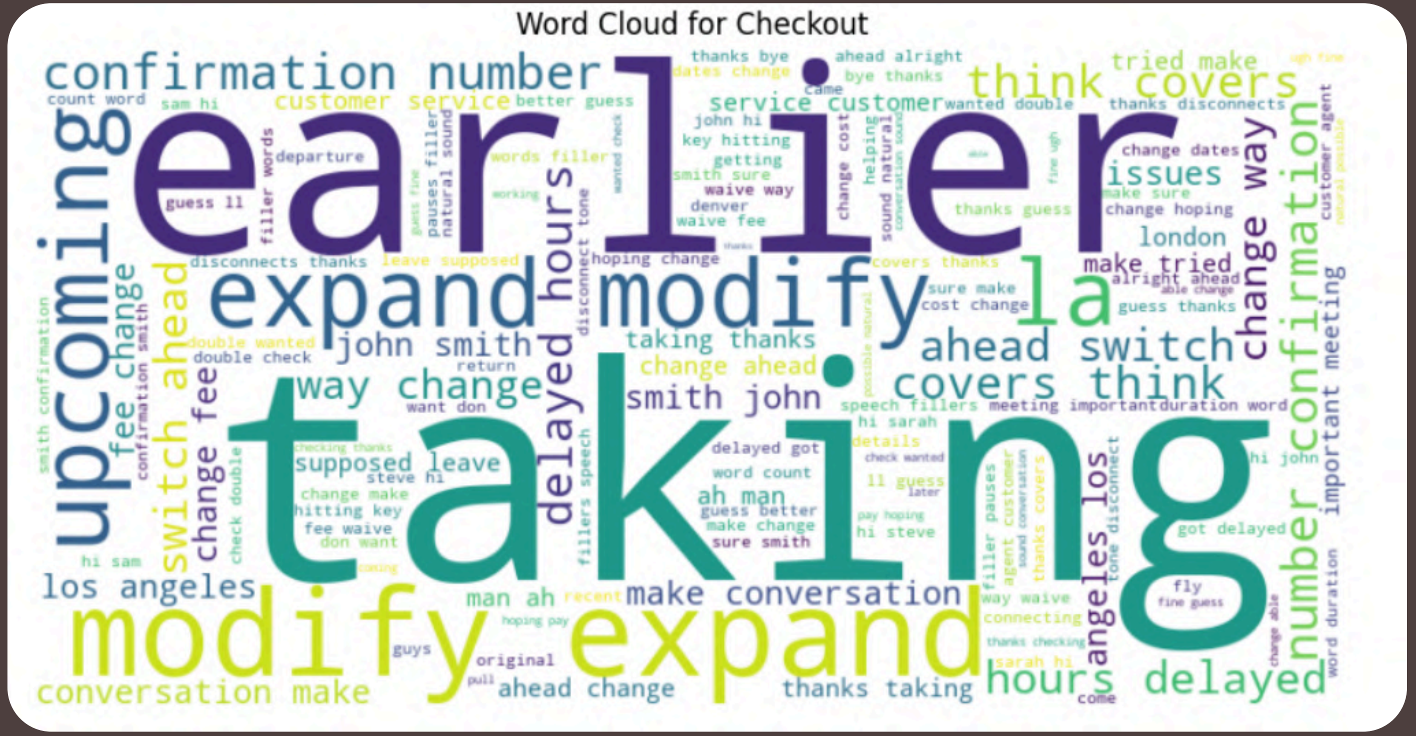
Approach:

- We made a dictionary of pre-defined IVR keywords
- For each transcript we checked the presence of these keyword to analyze the calls which could be easily handled by IVR systems and did not need an agent.

- "Money back" appeared 1670 times, making it the most frequent issue.
- "Refund request" appeared 187 times.
- Other issues like "baggage policy" (116 times), "lost baggage" (112 times), and "reschedule flight" (6 times) are also frequent but occur less often.

Second Approach:

- We generated the bi-grams for each call reason and analyzed the words which are more frequent than the others. This can help us identify the repetitive nature of the problems faced by customers' and hence it can be solved via IVR system.



- After fetching these bigrams, we can actually train a model which will learn to classify the words as either 'ASKING INFORMATION' or 'ASKING FOR A SOLUTION'.
- Now every time when customer will call the center, IVR will ask the customer to speak their reason for the call.
- This will be heard by the IVR system and then can be converted from speech-to-text using SOTA models.
- If the speech contains words that are asking for INFORMATION rather than SOLUTIONS then directing that call towards an agent can be easily avoided, by giving the IVR restricted access to the database.
- By doing this, IVR system will automatically verify the customer and will deliver the demanded information
- For e.g. 'delay hours' means customer needs information whereas 'change way' is the problem which will need manual solution by the agent.

IVR Suggestions

- We can create an IVR option of lost baggage, which will ask customer to type in their baggage_id and IVR will raise a ticket of it.
- Another option could be of reservation change, in which IVR will ask for booking_id, reschedule date, time and location, and from database it will either speak out all the available slots or can raise a ticket to solve by backend team.
- Another important option could be of the feedback on IVR asking for whether the customer found the IVR option helpful or not, which will help the developers to be in the constant feedback loop to make further changes.
- Inquiry about the Flight Status can also be served via IVR options. This can also be overcome by keeping the customers updated about the changes using Text messages/Mail.
- In the last we can offer to direct the call to an agent if the issues are not listed in the IVR options above. All the above steps will help in enhancing customer experience and optimizing the call center processes.

DELIVERABLES 3

- APPROACH 1: Classical ML
- APPROACH 2: BERT Model
- PREDICTING THE PRIMARY CALL REASONS CORRESPONDING TO TRANSCRIPTS USING THE APPROACH PROPOSED ABOVE

Classical ML approach for Predicting Primary Call Reasons

Approach 1:

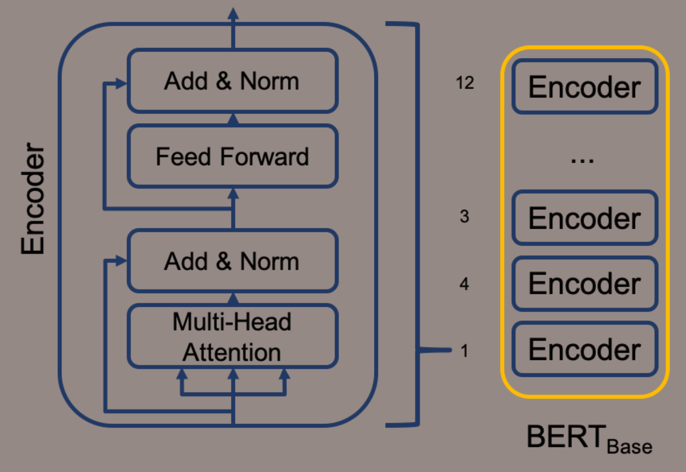
- To solve this multiclass classification problem we tried different classical machine learning models.
- We used same approach (as discussed earlier) to preprocess the data.
- In addition, we transformed the call transcript column to numerical feature using tf-idf method.
- We first created tf-idf matrices for each transcript and took mean of each, hence getting the number for each.
- Since the data was highly imbalance, we used a variant of SMOTE called SMOTEELM. This variant works good with muticlass problems.
- The AdaBoost classifier had low accuracy but it was able to predict 4 classes (others were not predicting more than 2).
- A better approach was needed to tackle the problem as we are predicting classes that had more frequency.

Transformer & Transfer learning approach for Predicting Primary Call Reasons

Approach 2:

- It is a multi-class classification problem which we solved by implementing BERT model.
- We passed the train dataset having columns: ‘cleaned_transcript’ & ‘primary_call_reason’ as the input.
- ‘primary_call_reason’ was label encoded to convert from categorical to numerical.
- Train-Val split of ratio 80:20 was made.
- We called Bert-Base-Uncased’s tokenizer and model from Hugging Face.
- This model was further trained on our train dataset for 2 epochs.

The Fine tuning we used



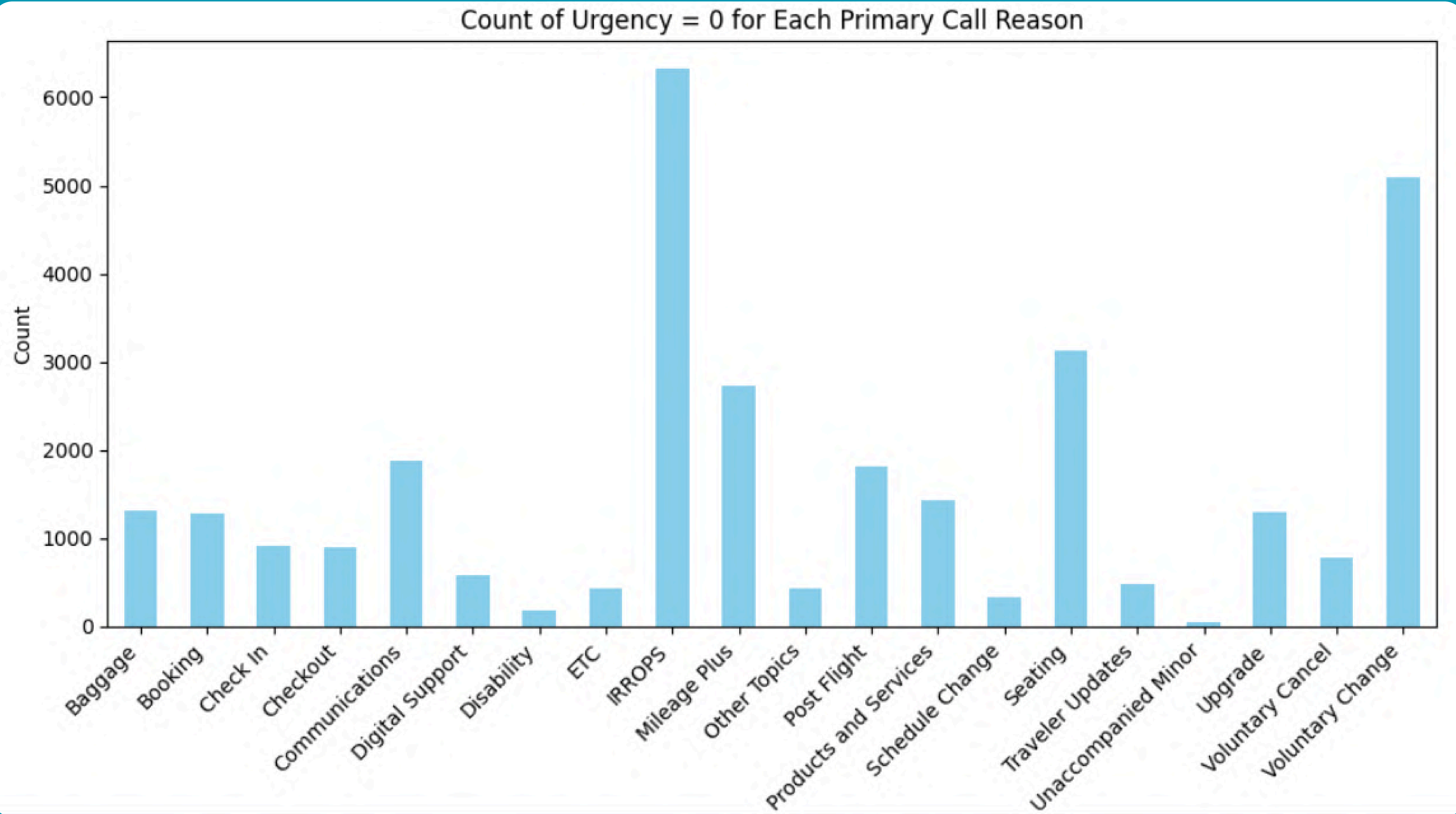
Bert base Model Architecture

- This fine-tuned model then was used to predict the Primary Call Reasons corresponding to each call id in the test dataset.
- The predicted reasons and their frequencies were:
    - IRROPS - 2368/5157 instances
    - Post Flight - 327/5157 instances
    - Voluntary Change - 2459/5157 instances
    - Mileage Plus - 3/5157

- We admit that our model's prediction was restricted to only 4 reasons, out of 20 in total. The root cause for this problem was the highly imbalanced nature of the train dataset.
- We could have associated the weights corresponding to minority reasons during computation of loss in training, but due to time & compute constraint, we were unable to proove that hypothesis.

BONUS: Reducing Waiting Time

- We analyzed the cleaned transcripts containing just customers’ side of conversation and classified the calls as URGENT or NOT_URGENT on basis of pre-defined keywords dictionary.
- The plot on next page shows the transcripts/calls that were classified as NOT URGENT corresponding to their primary call reason.
- So if we set up an IVR which will ask customer to press the button corresponding to the reason they called. we can easily queue the call based on their URGENCY.
- This will enhance the customer experience, solving the urgent problems earlier.


Count of Urgency = 0 for Each Primary Call Reason
Count
0
1000
2000
3000
4000
5000
6000
Baggage
Booking
Check In
Checkout
Communications
Digital Support
Disability
ETC
IRROPS
Mileage Plus
Other Topics
Post Flight
Products and Services
Schedule Change
Seating
Traveler Updates
Unaccompanied Minor
Upgrade
Voluntary Cancel
Voluntary Change

Just like United, our ideas are ready to soar – thanks for flying with us on this journey!