Decoding Emotional Landscapes in Friends' Interactions: Integrating Voice and Linguistic Analysis

Katia Djerroud- Cognitive computational neuroscience

BrainHack
School

Background

Theoretical Background:
The analysis of emotions in friends' interactions using voice and linguistic analysis is rooted in interdisciplinary research involving linguistics, psychology, computer science, and neuroscience.
Communication among friends includes rich verbal and non-verbal cues. Vocal features such as tone, rhythm, and pitch variations play a key role in emotional expression, while linguistic choices and sentence structures provide subtle hints about emotional states and interaction intensity.

Previous Works:
Voice Analysis: Tools like Praat show how acoustic features correlate with specific emotions.
Emotion Detection: OpenSmile extracts vocal features to identify emotions like joy, sadness, and anger.
Linguistic Analysis: NLTK and Hugging Face provide insights into how word choices and syntax reflect emotional states.
Integrative Models: Projects combining voice and text analysis, like those from SpeechBrain, create real-time emotion detection systems.

Literature:
Mehrabian, A. (1972): Highlights the importance of vocal and non-verbal cues in communication (Mehrabian's work).
Ekman, P., & Friesen, W. V. (1969): Influential research on facial expressions and emotions (Ekman's research).
Picard, R. W. (1997): "Affective Computing" explores how machines recognize and respond to human emotions (Affective Computing).

Main question

How can the integration of voice analysis, linguistic analysis, and natural language processing accurately detect and interpret emotions in interactions between friends, and what insights can this provide about the relational closeness between individuals?

Objectives

```
1  Develop and Validate a Comprehensive Model
2  Enhance Understanding of Emotional Expression
3  Explore Relationships Between Verbal and Non-Verbal Cues
4  Apply Advanced Tools and Techniques
5  Create a Comprehensive Dataset
6  Implement Real-Time Emotion Detection
7  Produce Interactive Visualizations
8  Contribute to Research in Human-Machine Interaction
9
```

Methods

```
 Collect Data
     |
     v
  Preprocess Data
     |
     v
Voice Analysis -> Extract Vocal Features (Praat, OpenSmile) -> Identify Emotional Cues
     |
     v
Linguistic Analysis -> Tokenization, Parsing, Sentiment Analysis (NLTK, Hugging Face) -> Detect Emotional States
     |
     v
Integrate Data -> Combine Vocal and Textual Features -> Model Training (SpeechBrain)
     |
     v
Emotion Detection Model -> Train and Validate Model -> Predict Emotions
     |
     v
Interactive Visualizations -> Generate Visual Representations -> Insights into Emotional States and Linguistic Patterns
```

Dataset

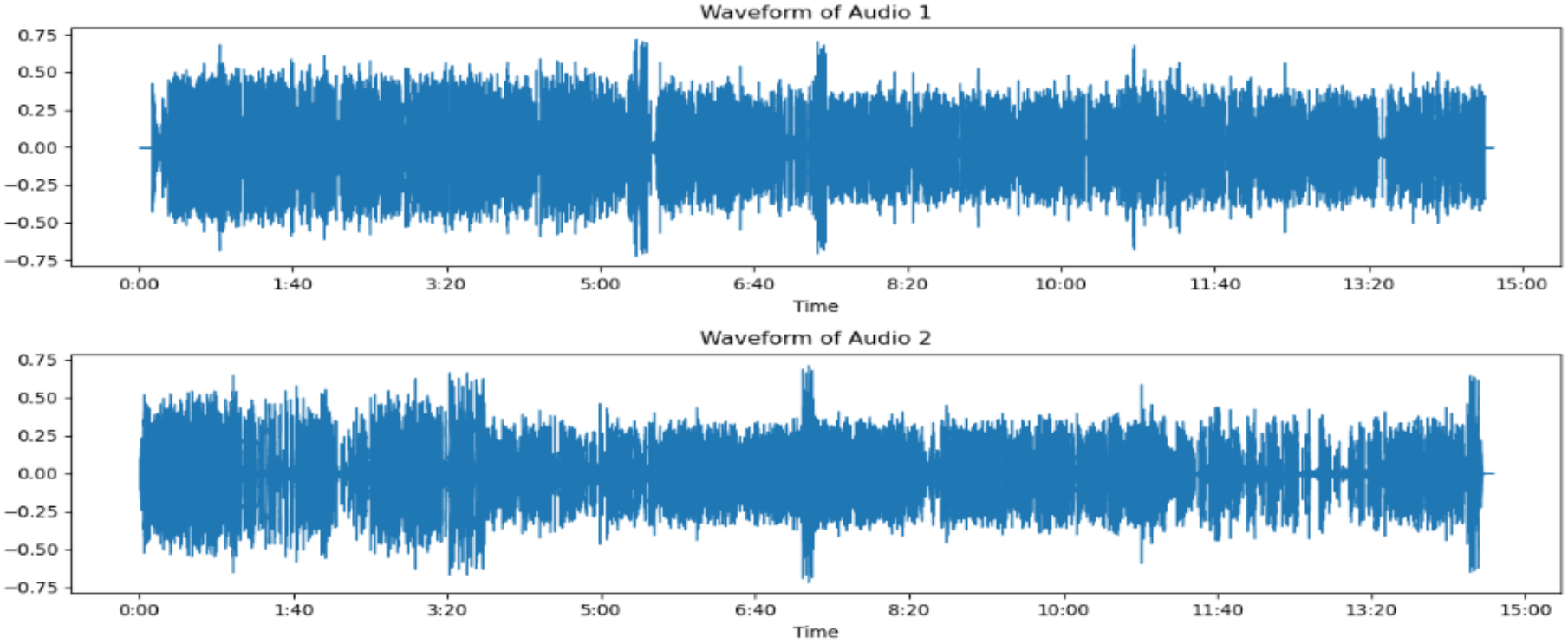
1/ Description: The dataset used for this project consists of episodes from the TV show "Friends", encompassing audio recordings, transcripts, and video footage. This dataset was obtained from MarieSTL, affiliated with Pierre Bellec's Lab.

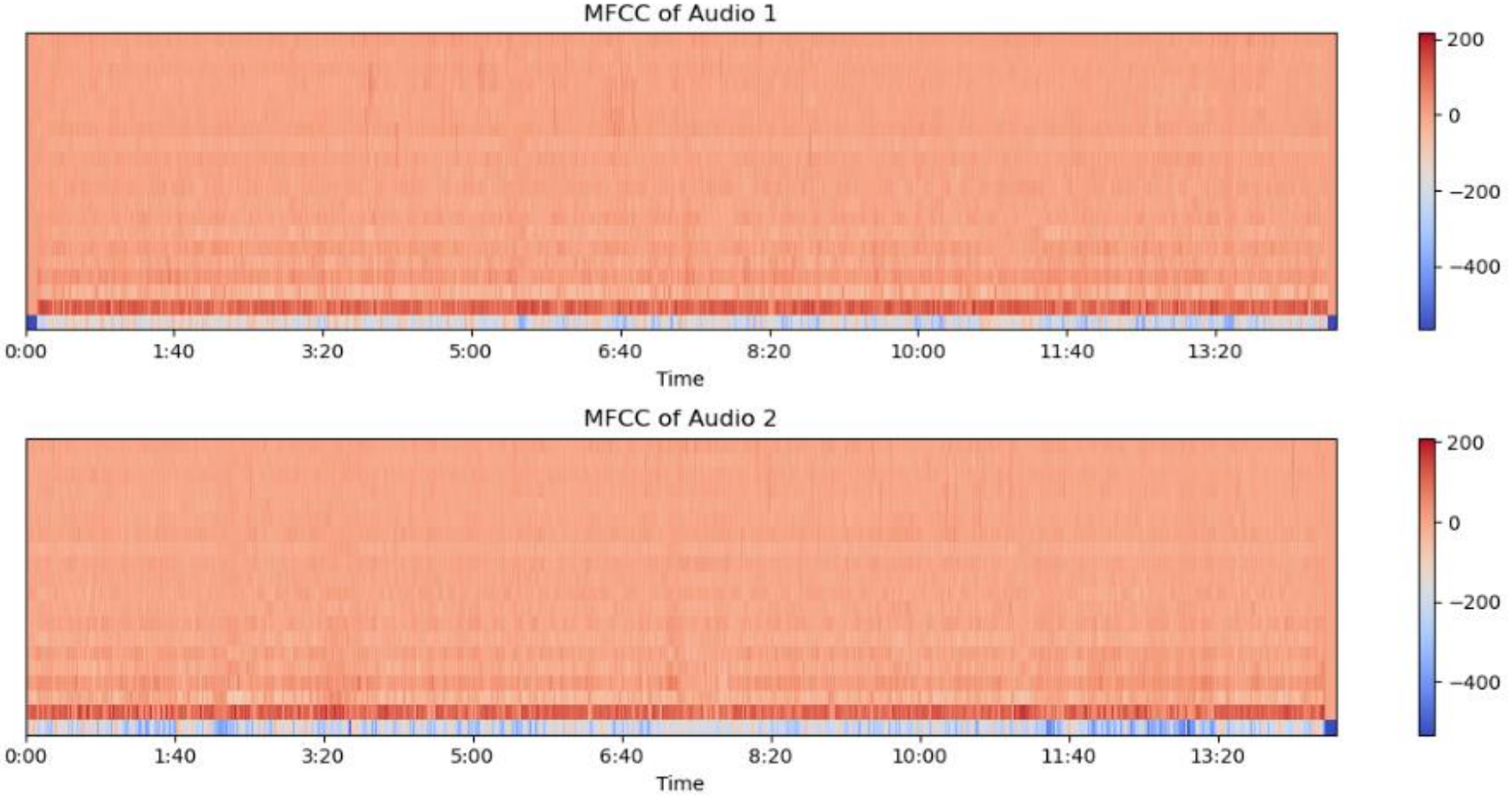
2/ Dataset Preparation:
Audio MP3 Files:

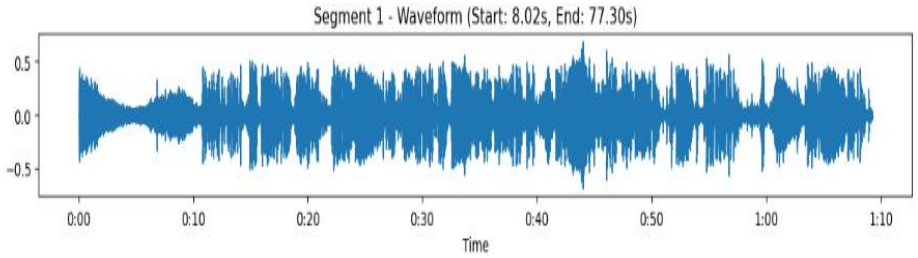
Format: MP3 audio files extracted from the video episodes of "Friends".
Content: These files contain the speech and vocal expressions of characters during interactions.
Usage: Used for voice analysis to extract acoustic features such as spectrograms, pitch, and formants using tools like Praat or Librosa.
JSON Transcript Files:
Format: JSON files containing textual transcripts of the dialogues from the episodes.
Content: Provides textual data of the conversations including dialogue lines, speaker information, and possibly timestamps.
Usage: Used for linguistic analysis, sentiment analysis, and text-based emotion detection. Tools like NLTK and Hugging Face Transformers can tokenize, parse, and analyze the sentiment and emotions expressed in the dialogues.

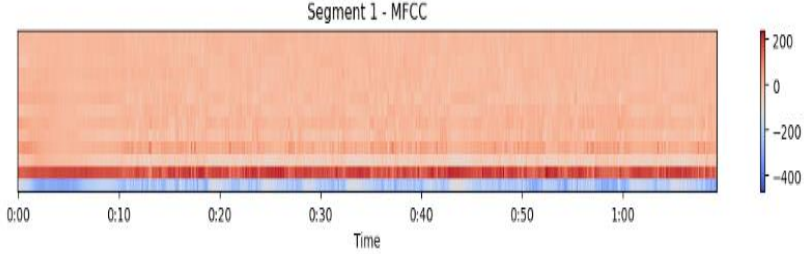
3/ Advantages of the Approach
Multi-Modal Analysis: By using both audio and textual data, you can conduct a comprehensive analysis of emotional dynamics in human interactions portrayed in "Friends".
Realism and Diversity: The dataset's origins from a popular TV show provide realistic and diverse scenarios, enriching the analysis with various emotional expressions and interpersonal dynamics.
Research Relevance: The conversion to audio MP3 and availability of transcript JSON files make the dataset suitable for answering your research question on discerning emotional dynamics using voice, linguistic, and text analysis techniques.

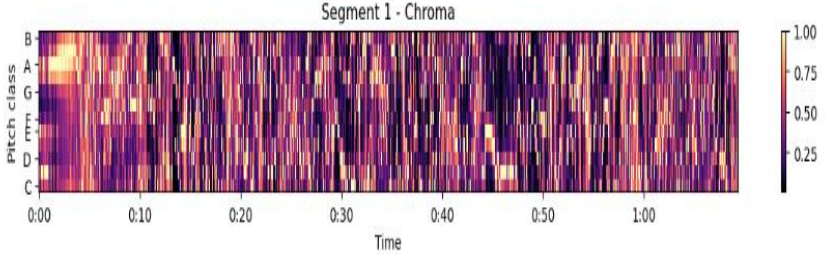
Results

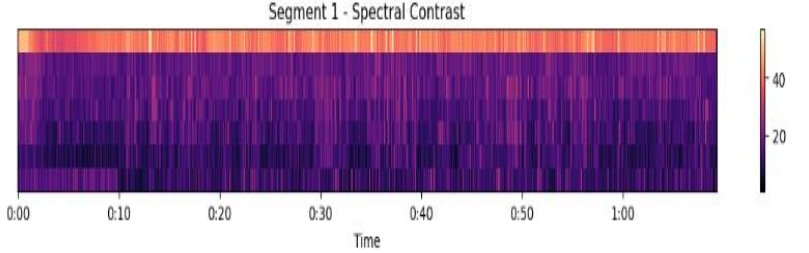

Waveform of Audio 1
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
0:00
1:40
3:20
5:00
6:40
8:20
10:00
11:40
13:20
15:00
Time
Waveform of Audio 2
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
0:00
1:40
3:20
5:00
6:40
8:20
10:00
11:40
13:20
15:00
Time

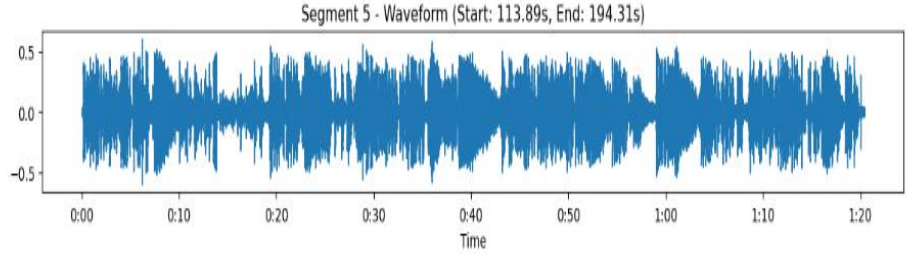

MFCC of Audio 1
200
0
−200
−400
0:00
1:40
3:20
5:00
6:40
8:20
10:00
11:40
13:20
Time
MFCC of Audio 2
200
0
−200
−400
0:00
1:40
3:20
5:00
6:40
8:20
10:00
11:40
13:20
Time

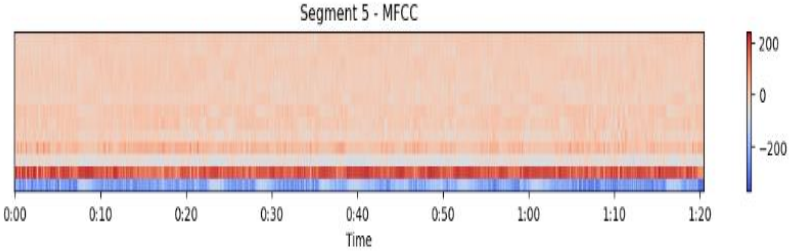

Segment 1 - Waveform (Start: 8.02s, End: 77.30s)
0.5
0.0
−0.5
0:00
0:10
0:20
0:30
0:40
0:50
1:00
1:10
Time

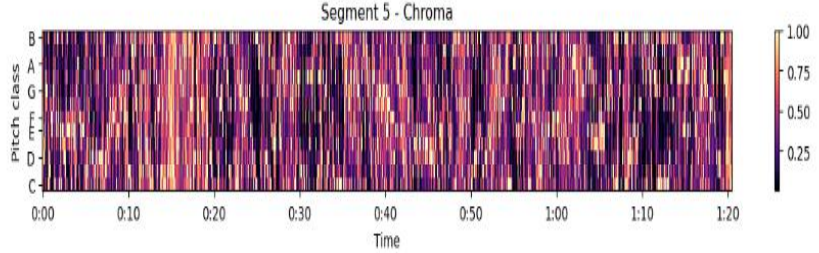

Segment 1 - MFCC
200
0
−200
−400
0:00
0:10
0:20
0:30
0:40
0:50
1:00
Time

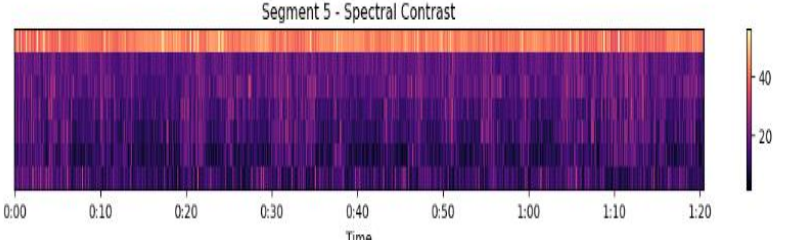

Segment 1 - Chroma
Pitch class
B
A
G
D
C
1.00
0.75
0.50
0.25
0:00
0:10
0:20
0:30
0:40
0:50
1:00
Time


Segment 1 - Spectral Contrast
40
20
0:00
0:10
0:20
0:30
0:40
0:50
1:00
Time

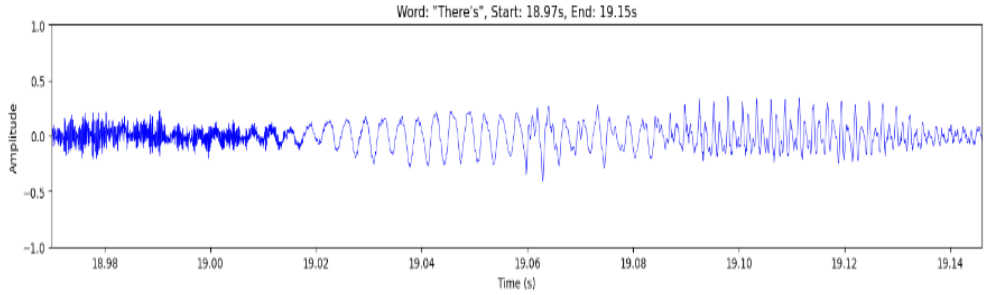

Segment 5 - Waveform (Start: 113.89s, End: 194.31s)
0.5
0.0
−0.5
0:00
0:10
0:20
0:30
0:40
0:50
1:00
1:10
1:20
Time

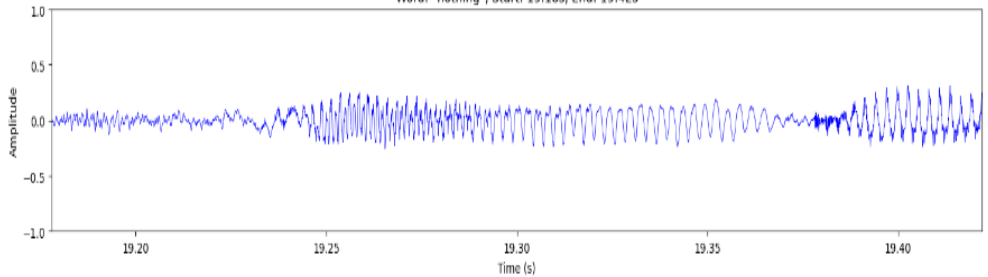

Segment 5 - MFCC
200
0
−200
0:00
0:10
0:20
0:30
0:40
0:50
1:00
1:10
1:20
Time

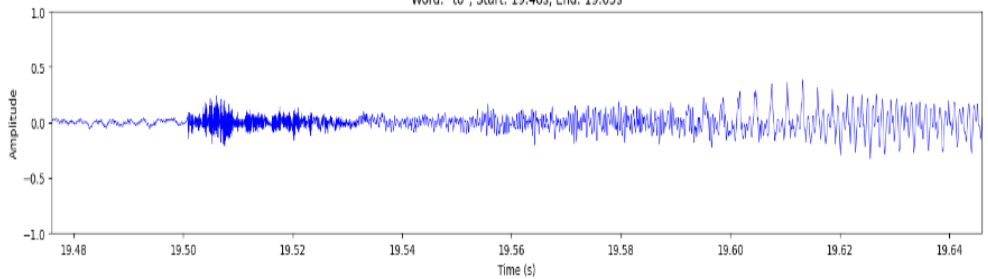

Segment 5 - Chroma
Pitch class
B
A
G
D
C
1.00
0.75
0.50
0.25
0:00
0:10
0:20
0:30
0:40
0:50
1:00
1:10
1:20
Time

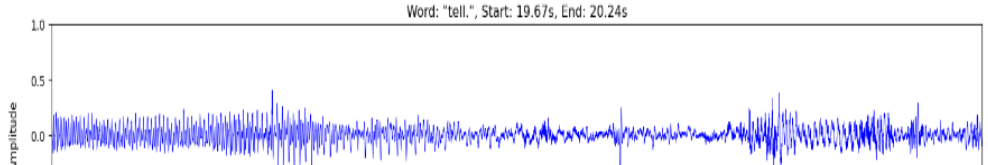

Segment 5 - Spectral Contrast
40
20
0:00
0:10
0:20
0:30
0:40
0:50
1:00
1:10
1:20
Time

```
[{'word': "There's", 'start': 18.97, 'end': 19.146, 'confidence': 0.6398},
 {'word': 'nothing', 'start': 19.178, 'end': 19.422, 'confidence': 0.99923},
 {'word': 'to', 'start': 19.476, 'end': 19.646, 'confidence': 1.0},
 {'word': 'tell.', 'start': 19.668, 'end': 20.24, 'confidence': 0.99896},
 {'word': 'Just', 'start': 20.85, 'end': 21.262, 'confidence': 0.5221},
 {'word': 'some', 'start': 21.316, 'end': 21.534, 'confidence': 0.99967},
 {'word': 'guy', 'start': 21.572, 'end': 21.726, 'confidence': 0.99993},
 {'word': 'I', 'start': 21.748, 'end': 21.886, 'confidence': 1.0},
 {'word': 'work', 'start': 21.908, 'end': 22.142, 'confidence': 0.77104},
 {'word': 'with.', 'start': 22.196, 'end': 22.75, 'confidence': 0.99895},
 {'word': 'Come', 'start': 22.9, 'end': 23.262, 'confidence': 0.99886},
 {'word': 'on,', 'start': 23.316, 'end': 23.774, 'confidence': 0.99997},
 {'word': "you're", 'start': 23.892, 'end': 24.314, 'confidence': 0.76648},
 {'word': 'going', 'start': 24.362, 'end': 24.526, 'confidence': 0.99987},
 {'word': 'out', 'start': 24.548, 'end': 24.734, 'confidence': 0.81},
 {'word': 'with', 'start': 24.772, 'end': 24.878, 'confidence': 0.99939},
 {'word': 'the', 'start': 24.884, 'end': 25.006, 'confidence': 0.93},
 {'word': 'guy.', 'start': 25.028, 'end': 25.214, 'confidence': 0.99937},
 {'word': "There's", 'start': 25.252, 'end': 25.418, 'confidence': 0.99932},
 {'word': 'got', 'start': 25.434, 'end': 25.518, 'confidence': 0.99832},
 {'word': 'to', 'start': 25.524, 'end': 25.646, 'confidence': 0.82},
 {'word': 'be', 'start': 25.668, 'end': 25.806, 'confidence': 0.99998},
 {'word': 'something', 'start': 25.828, 'end': 26.062, 'confidence': 0.99989
 {'word': 'wrong', 'start': 26.116, 'end': 26.334, 'confidence': 0.99996},
 {'word': 'with', 'start': 26.372, 'end': 26.526, 'confidence': 0.99996},
 {'word': 'him.', 'start': 26.548, 'end': 27.12, 'confidence': 0.92915},
 {'word': 'All', 'start': 27.81, 'end': 28.126, 'confidence': 0.9138},
 {'word': 'right,', 'start': 28.148, 'end': 28.286, 'confidence': 0.9999},
 {'word': 'joey,', 'start': 28.308, 'end': 28.666, 'confidence': 0.48141},
 {'word': 'be', 'start': 28.698, 'end': 28.894, 'confidence': 0.99876},
 {'word': 'nice.', 'start': 28.932, 'end': 29.52, 'confidence': 0.99594},
 {'word': 'So', 'start': 30.01, 'end': 30.374, 'confidence': 0.98573},
 {'word': 'does', 'start': 30.412, 'end': 30.566, 'confidence': 0.99985},
 {'word': 'he', 'start': 30.588, 'end': 30.678, 'confidence': 0.98085},
 {'word': 'have', 'start': 30.684, 'end': 30.806, 'confidence': 0.99997},
 {'word': 'a', 'start': 30.828, 'end': 30.918, 'confidence': 1.0},
 {'word': 'hump?', 'start': 30.924, 'end': 31.106, 'confidence': 0.97814},
 {'word': 'A', 'start': 31.138, 'end': 31.238, 'confidence': 1.0},
 {'word': 'hump', 'start': 31.244, 'end': 31.426, 'confidence': 0.84789},
 {'word': 'and', 'start': 31.458, 'end': 31.558, 'confidence': 1.0},
```







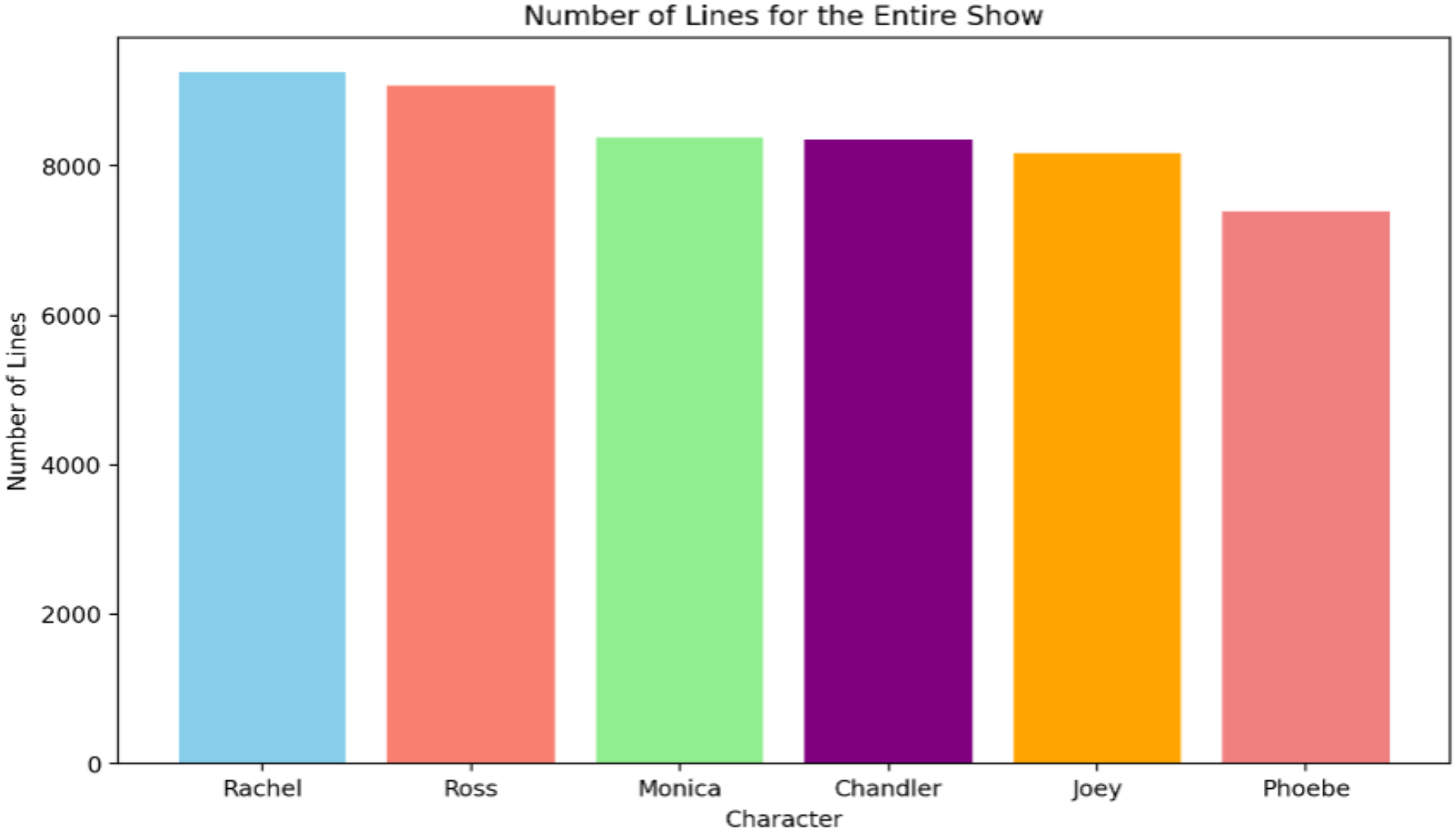
Transcription with Speaker Labels:
Speaker A: There's nothing to tell. It's just some guy I work with.
Speaker B: Come on. You're going out with the guy? There's got to be something wrong with him.
Speaker C: All right, Joey, be nice. So, does he have a hump? A hump and a hairpiece.
Speaker D: Wait, does he eat chalk? Just cause I don't want her to go through what I went through with Carl.
Speaker A: Oh, okay, everybody relax. Relax? This is not even a date. It's not? It's just two people going out to dinner and not having sex.
Speaker C: Sounds like a date to me. All right, so I'm back in high school. I'm standing in the middle of the cafeteria, and I realize I am totally naked. Then I look down, and I realize there is a phone there.
Speaker B: Instead of.
Speaker C: That's right. All of a sudden, the phone starts to ring. Now I don't know what to do. Everybody starts looking at me.
Speaker A: They weren't looking at you before?
Speaker C: Finally, I figure I'd better answer it turns out it's my mother. Which is very, very weird because she never calls me.
Speaker E: Hi.
Speaker B: This guy says, hello, I want to kill myself.
Speaker A: You okay, sweetie?
Speaker F: I just feel like someone reached down my throat, grabbed my small intestine, pulled it out of my mouth and tied it around my neck.
Speaker C: Cookie?
Speaker A: Carol moved her stuff out today. Let me get you some coffee.
Speaker F: Thanks. Ooh, no. Oh, no, no. Don't. Stop cleansing my aura. Don't. Just leave my aura alone.
Speaker D: Okay, fine. Be murky.
Speaker F: I'll be fine. All right. Really, everyone? I hope she'll be very happy.
Speaker E: No, you don't.
Speaker F: No, I don't. To hell with her. She left me.
Speaker B: And you never knew she was a lesbian?
Speaker F: No. Okay, why does everyone keep fixating on that? She didn't know. How should I know?
Speaker C: Sometimes I wish I was a lesbian. Did I say that out loud?
Speaker F: I told mom and dad last night. They. They seem to take it pretty well.
Speaker A: Oh, really? So that hysterical phone call I got from a woman sobbing at 03:00 a.m. i'll never have grandchildren? I'll never have grandchildren. Was what? A wrong number?
Speaker F: Sorry.
Speaker B: All right, Ross, look, you're feeling a lot of pain right now. You're angry, you're hurtin'can. I tell you what the answer is? Strip joints. Come on, you're single.
Speaker F: Have some hormones. See, but I don't want to be single, okay?

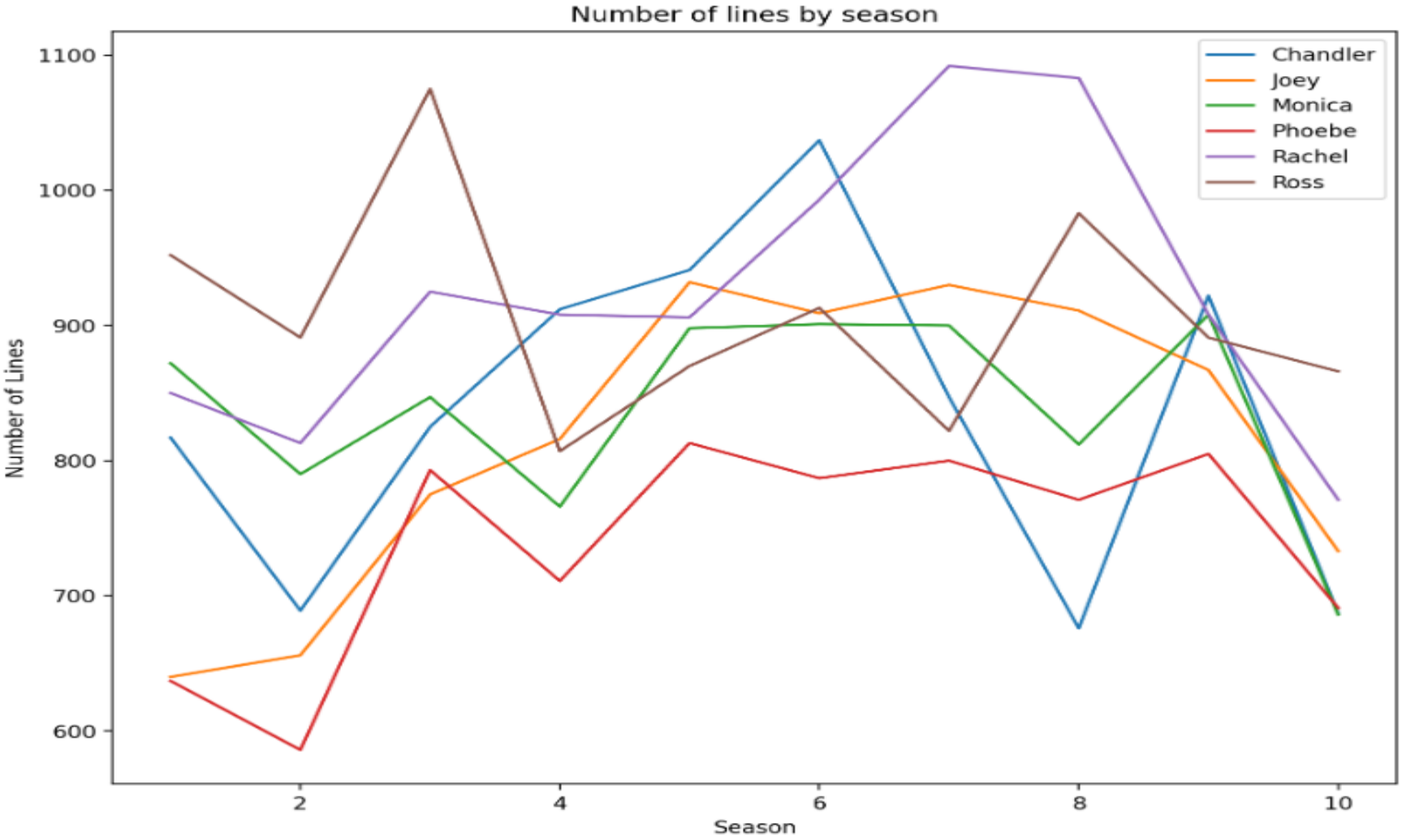
Entities Detected:
Type: person_name, Text: Joey
Type: person_name, Text: Carl
Type: person_name, Text: Cookie
Type: person_name, Text: Carol
Type: gender_sexuality, Text: lesbian
Type: gender_sexuality, Text: lesbian
Type: time, Text: 03:00 a.m
Type: person_name, Text: Ross
Type: marital_status, Text: single
Type: marital_status, Text: single
Type: marital_status, Text: married
Type: money_amount, Text: million dollar
Type: person_name, Text: Rachel
Type: person_name, Text: Monica
Type: person_name, Text: Rachel
Type: organization, Text: Lincoln high
Type: person_name, Text: Chandler
Type: person_name, Text: Phoebe
Type: person_name, Text: Joey
Type: person_name, Text: Ross
Type: duration, Text: a half hour
Type: person_name, Text: Barry
Type: person_name, Text: Barry
Type: person_name, Text: potato head
Type: person_name, Text: Christine
Type: person_name, Text: Monica
Type: person_name, Text: Monica
Type: person_name, Text: Joey
Type: person_name, Text: Chandler
Type: person_name, Text: Joey
Type: person_name, Text: Paul
Type: time, Text: 630
Type: person_name, Text: Paul
Type: person_name, Text: Paul
Type: occupation, Text: wine guy
Type: person_name, Text: Paul
Type: person_name, Text: Paul

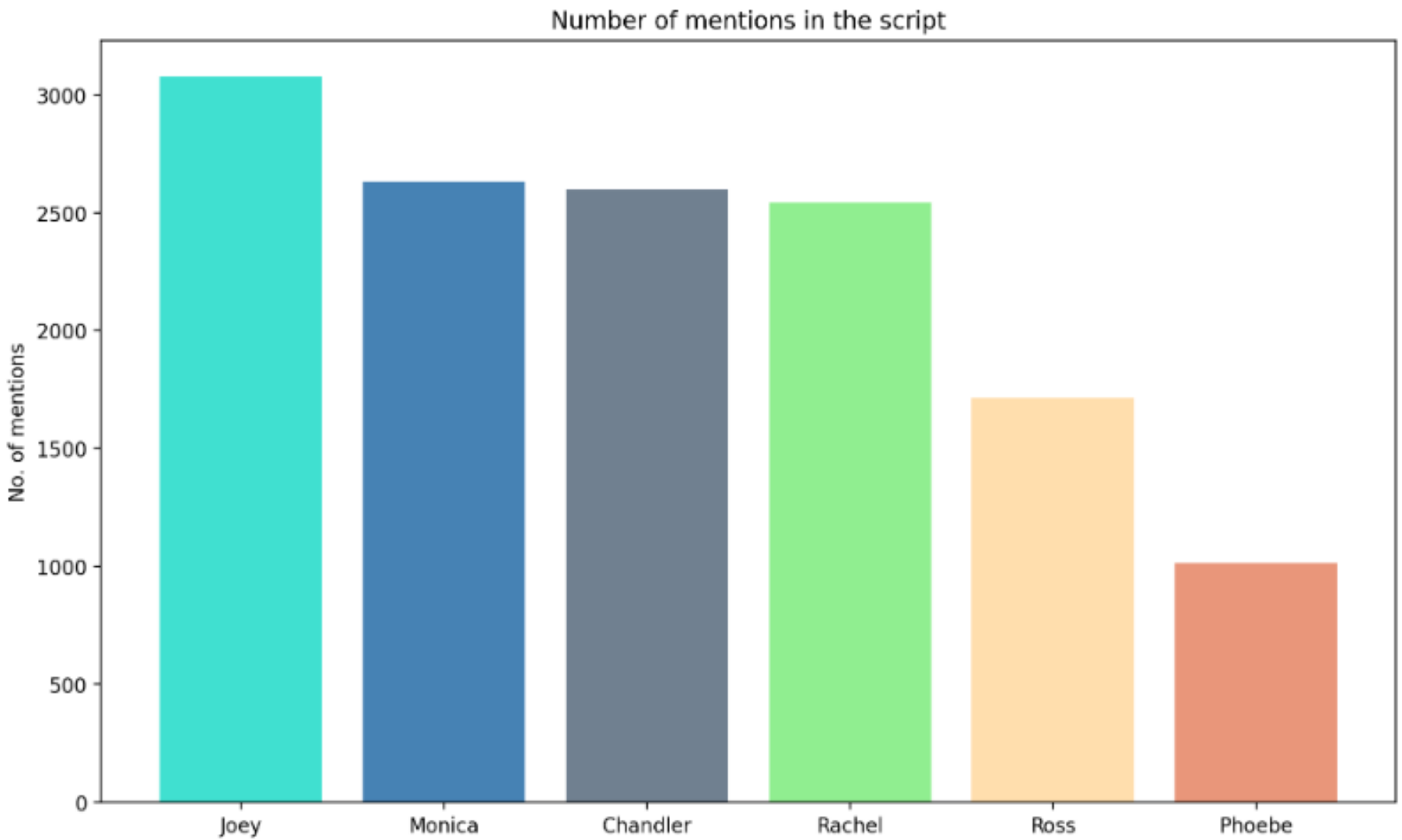
```
Sentence: 'There's nothing to tell.'
Sentiment: NEGATIVE, Score: 0.9991669654846191
Sentence: 'Just some guy I work with.'
Sentiment: POSITIVE, Score: 0.9709184169769287
Sentence: 'Come on, you're going out with the guy.'
Sentiment: POSITIVE, Score: 0.9983578324317932
Sentence: 'There's got to be something wrong with him.'
Sentiment: NEGATIVE, Score: 0.9969286322593689
Sentence: 'All right, joey, be nice.'
Sentiment: POSITIVE, Score: 0.9997671246528625
Sentence: 'So does he have a hump?'
Sentiment: NEGATIVE, Score: 0.9930300116539001
Sentence: 'A hump and a hairpiece.'
Sentiment: NEGATIVE, Score: 0.782961368560791
Sentence: 'Wait, does he eat chalk?'
Sentiment: NEGATIVE, Score: 0.9989175796508789
Sentence: 'Just because I don't want her to go through what I went through with Carl okay, everybody relax.'
Sentiment: POSITIVE, Score: 0.9936161041259766
Sentence: 'Relax.'
Sentiment: POSITIVE, Score: 0.997666597366333
Sentence: 'This is not even a date.'
Sentiment: NEGATIVE, Score: 0.9990111589431763
Sentence: 'It's not?'
Sentiment: NEGATIVE, Score: 0.9972633123397827
Sentence: 'It's just two people going out to dinner and not having sex.'
Sentiment: NEGATIVE, Score: 0.9959657192230225
Sentence: 'Sounds like a date to me.'
Sentiment: POSITIVE, Score: 0.9992936849594116
Sentence: 'All right, so I'm back in high school.'
Sentiment: POSITIVE, Score: 0.9988865256309509
Sentence: 'I'm standing in the middle of the cafeteria and I realize I am totally naked.'
Sentiment: NEGATIVE, Score: 0.7636355757713318
Sentence: 'Then I look down and I realize there is a phone there.'
Sentiment: NEGATIVE, Score: 0.9819121360778809
Sentence: 'Instead of that's right?'
Sentiment: NEGATIVE, Score: 0.9900956153869629
Sentence: 'All of a sudden, the phone starts to ring.'
```

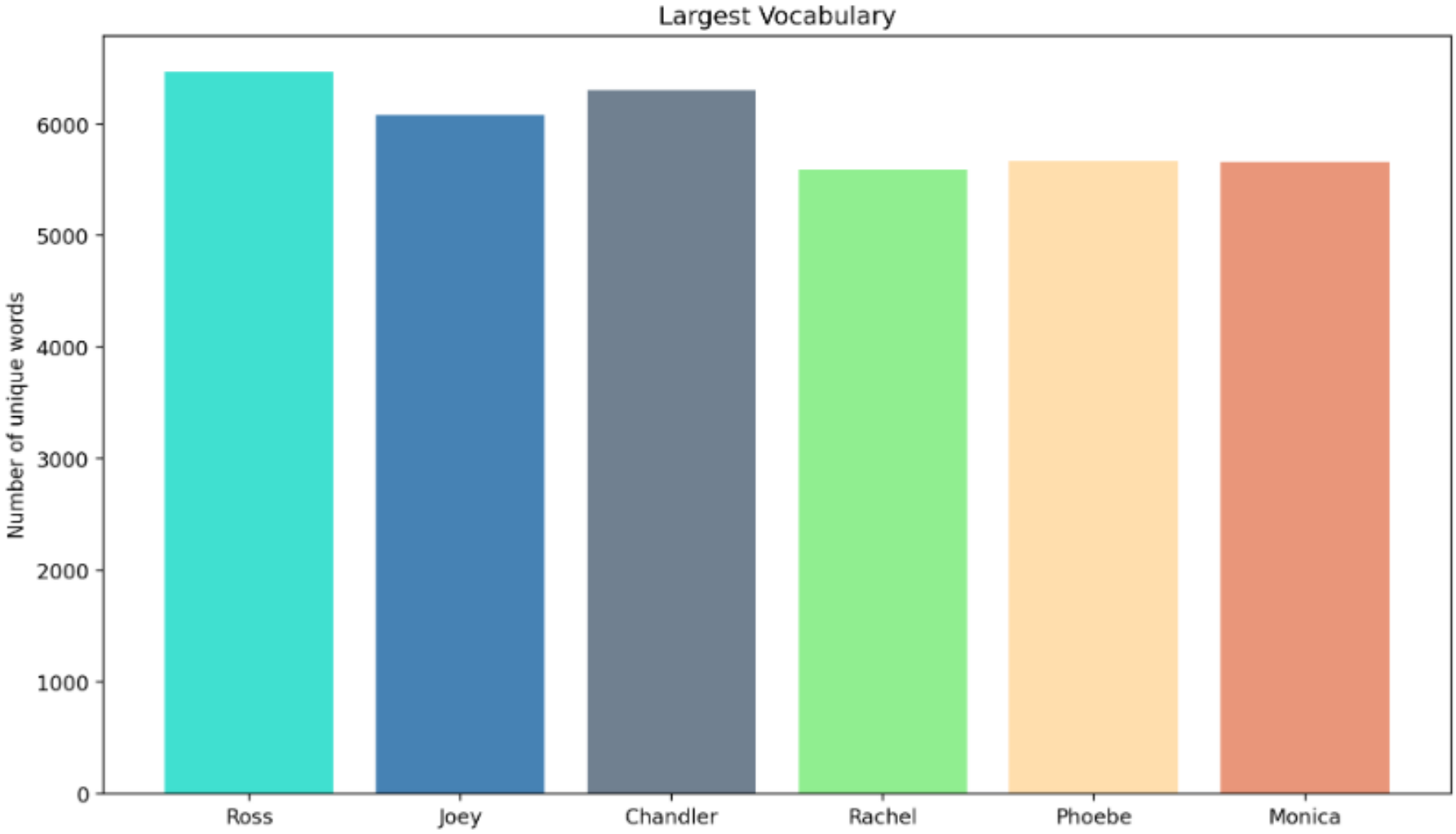
	season	episode	char	line
1	1	1	Monica	There's nothing to tell! He's just some guy
2	1	1	Joey	C'mon, you're going out with the guy! There's
3	1	1	Chandler	All right Joey, be
4	1	1	Phoebe	Wait, does he eat chalk?
5	1	1	Phoebe	Just, 'cause, I don't want her to go through
...	...	...	...	...
61755	9	9	Ross	(holds up the message) Uh, Rach.
61756	9	9	Rachel	(still looking at Emma) Yeah?
61757	9	9	Ross	(pauses, then crunches up the note and stuffs ...
61758	9	9	Ross	Hey you guys. (He sits.)
61760	9	9	Phoebe	Hey. I'll be right back. I've got to go to t...

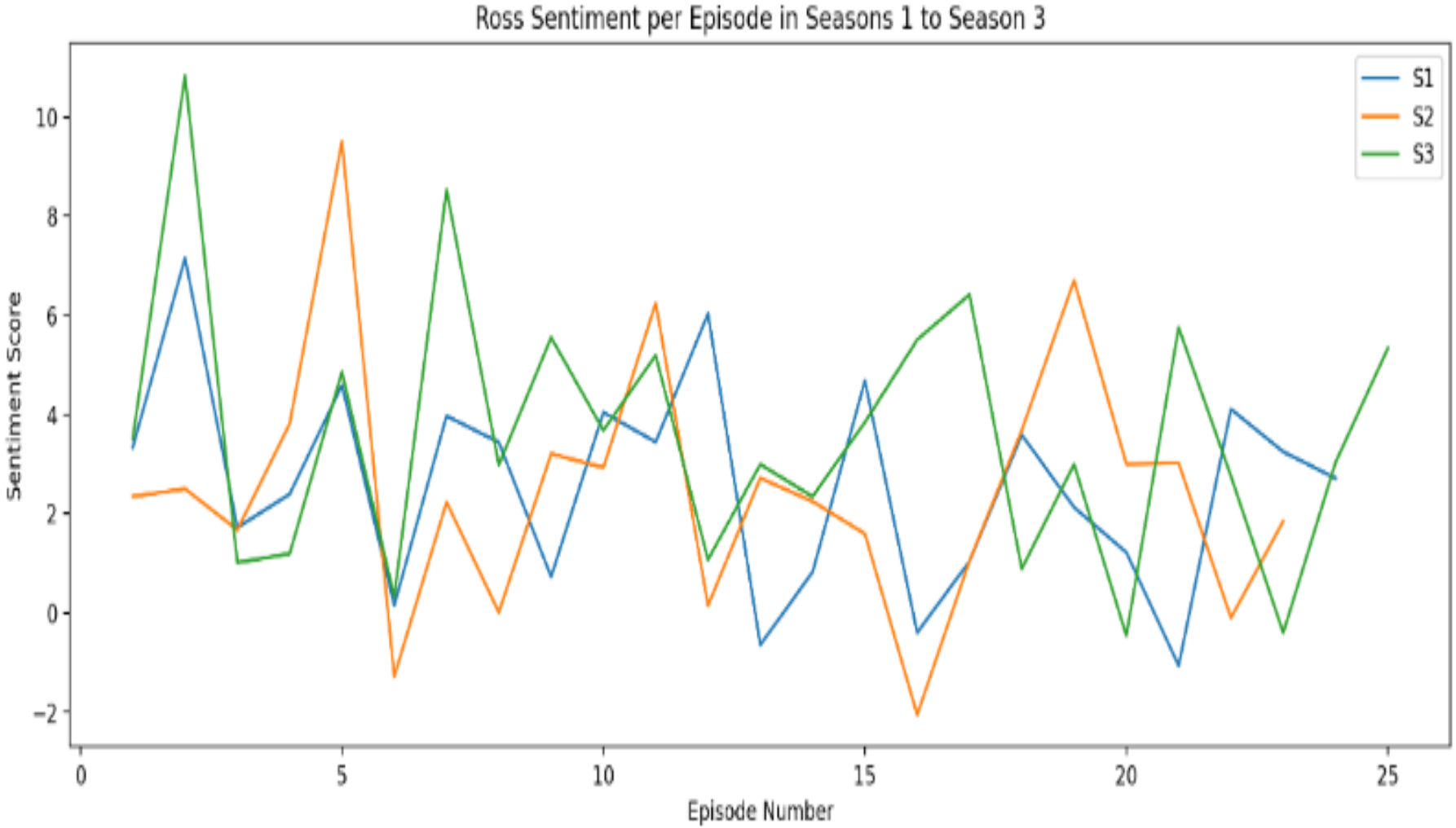
50615 rows × 4 columns


Number of Lines for the Entire Show
Number of Lines
0
2000
4000
6000
8000
Rachel
Ross
Monica
Chandler
Joey
Phoebe
Character


Number of lines by season
Number of Lines
Season
1100
1000
900
800
700
600
2
4
6
8
10
Chandler
Joey
Monica
Phoebe
Rachel
Ross


Number of mentions in the script
No. of mentions
3000
2500
2000
1500
1000
500
0
Joey
Monica
Chandler
Rachel
Ross
Phoebe


Largest Vocabulary
Number of unique words
0
1000
2000
3000
4000
5000
6000
Ross
Joey
Chandler
Rachel
Phoebe
Monica


Ross Sentiment per Episode in Seasons 1 to Season 3
S1
S2
S3
Sentiment Score
Episode Number
10
8
6
4
2
0
−2
0
5
10
15
20
25

	Character 1	Character 2	Shared Vocabulary	Percentage
0	Chandler	Joey	[All, right, Joey,, be, Sounds, like, a, date,...	93.786976
1	Chandler	Monica	[All, right, Joey,, be, Sounds, like, a, date,...	92.429344
2	Chandler	Phoebe	[All, right, Joey,, be, Sounds, like, a, date,...	95.830516
3	Chandler	Rachel	[All, right, Joey,, be, like, a, date, to, me....	89.925757
4	Chandler	Ross	[All, right, Joey,, be, Sounds, like, a, date,...	95.229853
5	Joey	Chandler	[C'mon,, you're, going, out, with, the, guy!, ...	93.786976
6	Joey	Monica	[C'mon,, you're, going, out, with, the, guy!, ...	91.100023
7	Joey	Phoebe	[C'mon,, you're, going, out, with, the, guy!, ...	97.836376
8	Joey	Rachel	[C'mon,, you're, going, out, with, the, guy!, ...	89.741038
9	Joey	Ross	[C'mon,, you're, going, out, with, the, guy!, ...	94.853631
10	Monica	Chandler	[There's, nothing, to, He's, just, some, guy, ...	92.429344
11	Monica	Joey	[There's, nothing, to, tell!, He's, just, some...	
12	Monica	Phoebe	[There's, nothing, to, He's, just, some, guy, ...	
13	Monica	Rachel	[There's, nothing, to, He's, just, some, guy, ...	
14	Monica	Ross	[There's, nothing, to, He's, just, some, guy, ...	
15	Phoebe	Chandler	[Wait,, does, he, eat, 'cause,, I, don't, want...	
16	Phoebe	Joey	[Wait,, does, he, eat, Just,, I, don't, want, ...	
17	Phoebe	Monica	[Wait,, does, he, eat, Just,, I, don't, want, ...	
18	Phoebe	Rachel	[Wait,, does, he, eat, Just,, I, don't, want, ...	
19	Phoebe	Ross	[Wait,, does, he, eat, Just,, 'cause,, I, don'...	
20	Rachel	Chandler	[Oh, God, Monica, hi!, Thank, God!, I, just, w...	
21	Rachel	Joey	[Oh, God, Monica, hi!, Thank, God!, I, just, w...	
22	Rachel	Monica	[Oh, God, Monica, hi!, Thank, God!, I, just, w...	
23	Rachel	Phoebe	[Oh, God, Monica, hi!, Thank, God!, I, just, w...	
24	Rachel	Ross	[Oh, God, Monica, hi!, Thank, God!, I, just, w...	
25	Ross	Chandler	[Hi., I, just, feel, like, someone, reached, d...	
26	Ross	Joey	[Hi., I, just, feel, like, someone, down, my, ...	
27	Ross	Monica	[Hi., I, just, feel, like, someone, down, my, ...	
28	Ross	Phoebe	[Hi., I, just, feel, like, someone, down, my, ...	
29	Ross	Rachel	[Hi., I, just, feel, like, someone, down, my, ...	

Number of lines spoken by each character across all seasons:

```
char
Rachel     9249
Ross       9070
Monica     8380
Chandler   8353
Joey       8169
Phoebe     7394
Name: count, dtype: int64
```

Total number of lines spoken across all seasons: 50615

Proportion of lines spoken by each character across all seasons:

```
char
Rachel     0.182732
Ross       0.179196
Monica     0.165564
Chandler   0.165030
Joey       0.161395
Phoebe     0.146083
Name: count, dtype: float64
```

Participation is balanced across all seasons.

```
Chandler vs Joey: LSM Score = 0.98
Chandler vs Monica: LSM Score = 0.99
Chandler vs Phoebe: LSM Score = 0.98
Chandler vs Rachel: LSM Score = 0.97
Chandler vs Ross: LSM Score = 0.99
Joey vs Monica: LSM Score = 0.97
Joey vs Phoebe: LSM Score = 0.97
Joey vs Rachel: LSM Score = 0.96
Joey vs Ross: LSM Score = 0.99
Monica vs Phoebe: LSM Score = 0.98
Monica vs Rachel: LSM Score = 0.98
Monica vs Ross: LSM Score = 0.98
Phoebe vs Rachel: LSM Score = 0.99
Phoebe vs Ross: LSM Score = 0.98
Rachel vs Ross: LSM Score = 0.97
```

```
Turn Counts:
Rachel: 3 turns
Ross: 1 turns
Monica: 2 turns
Chandler: 1 turns
Joey: 1 turns
Phoebe: 1 turns

Interruptions: 8
```

Frequency of Interaction:

	char	interaction_count
0	Chandler	8353
1	Joey	8169
2	Monica	8380
3	Phoebe	7394
4	Rachel	9249
5	Ross	9070

	char	line	keywords
0	Rachel	I love shopping!	[(love, 1), (shopping!, 1)]
1	Ross	How are you?	[(how, 1), (you?, 1)]
2	Monica	Let's cook dinner.	[(let's, 1), (cook, 1), (dinner., 1)]
3	Chandler	Could you pass the remote?	[(could, 1), (pass, 1), (remote?, 1)]
4	Joey	How you doin?	[(how, 1), (doin?, 1)]
5	Phoebe	I wrote a new song.	[(wrote, 1), (new, 1), (song., 1)]

Deliverables

•**GitHub Repository:**
•Repository hosting all project-related materials including code scripts, notebooks, and documentation.

•**Code Scripts and Jupyter Notebooks:**
•Python scripts and Jupyter notebooks implementing algorithms and analysis pipelines for voice analysis, linguistic analysis, emotion detection, etc.

•**Documentation:**
•README File: Overview of the project, installation instructions, and how to use the code.
•Project Documentation: Detailed description of the methodology, algorithms used, and interpretation of results.

•**Dataset(s):**
•Data used for analysis (audio MP3 files, JSON transcript files from "Friends" episodes).

•**Analysis Workflow/Pipeline:**
•Diagram or document outlining the workflow from data collection and preprocessing to analysis and visualization.

•**Training Materials:**
•Scripts, configuration files, and models developed for emotion detection and other analyses.

•**Reports and Presentations:**
•Summary reports of analysis results, insights derived from the data, and findings related to emotional dynamics in human interactions.

Conclusion

Achievements

1. **Integration of Voice and Linguistic Analysis:** Successfully implemented algorithms for voice analysis using Librosa, Praat, and OpenSMILE to extract acoustic features and analyze vocal characteristics. Combined with NLTK, SpaCy, and Hugging Face Transformers for linguistic analysis, allowing for comprehensive exploration of emotional dynamics in social interactions based on both speech and text data.
2. **Emotion Detection Model Development:** Developed and fine-tuned emotion detection models using Hugging Face Transformers and OpenSMILE, enabling accurate identification of emotional states from audio and textual inputs.
3. **Analysis Pipeline Implementation:** Established a robust analysis pipeline from data preprocessing to visualization, facilitating systematic exploration of the "Friends" dataset to uncover patterns in emotional expression and linguistic behavior.
4. **Documentation and Sharing:** Created a well-documented GitHub repository containing code scripts, Jupyter notebooks, and comprehensive documentation (README, project documentation) to ensure reproducibility and share insights with the research community.

Data Integration Complexity: Managing and integrating multi-modal data (audio, transcripts) from "Friends" episodes posed challenges in data preprocessing and alignment between voice and text features.

Model Optimization: Fine-tuning emotion detection models required experimentation with hyperparameters and feature extraction methods to achieve optimal performance across different emotional states and contexts.

Tool Familiarization: Learning to effectively use and integrate diverse tools such as Praat for voice analysis and transformer-based models from Hugging Face for text analysis required time and effort.

Challenges Encountered

Tools Learned to Use

- **Librosa:** Python library for audio and music analysis.
- **Praat:** Software for phonetic analysis of speech.
- **OpenSMILE:** Feature extraction tool for audio data.
- **NLTK (Natural Language Toolkit):** Library for natural language processing.
- **SpaCy:** NLP library for tokenization, NER, and syntactic analysis.
- **Hugging Face Transformers:** Framework for training and using transformer-based models for NLP tasks
- **GitHub:** Version control and collaboration platform.

Next Steps

Enhance Model Performance: Continuously optimize emotion detection models for improved accuracy and robustness across diverse datasets and contexts. + combining with the video

Expand Dataset Coverage: Incorporate additional datasets beyond "Friends" episodes or other database to generalize findings and validate models in broader social interaction contexts.

Advanced Visualization Techniques: Explore advanced visualization methods to provide deeper insights into emotional dynamics and linguistic patterns.

Publication and Sharing: Prepare research findings for publication in academic journals for my phd and present insights at conferences to contribute to the field of computational cognitive neuroscience and human-computer interaction.

Collaboration and Feedback: Collaborate with peers and seek feedback from experts to refine methodologies and broaden the impact of research outcomes in understanding human emotional expression and communication.

References

Bibliography

1.Boersma, P., & Weenink, D. (2022). Praat: Doing Phonetics by Computer [Computer program]. Version 6.1.37. Retrieved from http://www.praat.org/

2.Eyben, F., Wöllmer, M., & Schuller, B. (2010). Opensmile: the Munich versatile and fast open-source audio feature extractor. In Proceedings of the international conference on multimedia (pp. 1459-1462).

3.Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. O'Reilly Media.

4.Manning, C. D., & Schütze, H. (1999). Foundations of Statistical Natural Language Processing. MIT Press.

5.Hugging Face Transformers: State-of-the-art Natural Language Processing for PyTorch and TensorFlow. Retrieved from https://huggingface.co/transformers/

6.Gensim: Topic Modelling for Humans. Retrieved from https://radimrehurek.com/gensim/

7.SpaCy: Industrial-strength Natural Language Processing in Python. Retrieved from https://spacy.io/

GitHub Repositories

1.Example GitHub Repository for NLP and Emotional Analysis:

1. Repository: ExampleNLPProject
2. Description: A comprehensive repository demonstrating NLP techniques and emotional analysis methodologies using Python and various libraries.

2.Emotion Detection and Voice Analysis Repository:

1. Repository: VoiceEmotionDetection
2. Description: Focuses on voice analysis and emotion detection using Praat, OpenSMILE, and machine learning techniques for acoustic feature extraction and emotional state recognition.

3.Friends TV Show Dataset Repository (Example):

1. Repository: FriendsDataset
2. Description: An example repository showcasing data preprocessing scripts, audio conversion tools, and JSON transcript handling for analyzing social interactions and emotional dynamics in "Friends" episodes.

Articles and Research Papers

1.Schuller, B., & Batliner, A. (2013). Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing. Wiley.

2.Jurafsky, D., & Martin, J. H. (2019). Speech and Language Processing (3rd ed.). Pearson.

3.Crystal, D. (2008). A Dictionary of Linguistics and Phonetics (6th ed.). Wiley-Blackwell.

Additional Resources

•ELAN: Linguistic Annotation Tool for annotating and analyzing linguistic data. More information at https://tla.mpi.nl/tools/tla-tools/elan/

•Inception: Tool for qualitative data analysis in linguistics. More information at https://inception-project.github.io/

Thank you for your Time