Analyzing WhatsApp Communication Trends

Introduction

Digital communication has transformed the way we interact, with messaging applications like WhatsApp leading the charge. WhatsApp, a widely-used platform, allows users to send text messages, voice notes, images, and videos, facilitating real-time communication across the globe. As of late 2024, the app boasts over 2 billion users, making it a rich source for analyzing digital interactions. This study delves into a dataset extracted from WhatsApp conversations, focusing on various dimensions of communication.

The dataset comprises over 62,000 messages exchanged between participants over a six-month period. This extensive collection allows for a comprehensive analysis of communication trends. The primary metrics considered include the total message count, which stands at approximately 35,000. Such a dataset provides a robust foundation for exploring how individuals convey their thoughts and emotions through digital means.

Key aspects of the analysis include temporal patterns in messaging, which examine how communication frequency varies throughout the day and week. Additionally, emotional expression is scrutinized to understand how participants convey feelings through text—whether they lean towards positivity, negativity, or neutrality. Topic clustering is another focal point, as it explores common themes and subjects that arise in conversations, revealing the interests and concerns of the participants.

Furthermore, sentiment distribution offers insights into the overall emotional tone of the conversations, allowing us to gauge the atmosphere of interactions. Lastly, a comparative analysis between participants will highlight differences in communication styles and emotional expression, contributing to a deeper understanding of interpersonal dynamics in a digital context. Through this multifaceted approach, the study aims to illuminate the intricate patterns and emotional landscapes inherent in WhatsApp communication.

Emoji Usage in Digital Communication

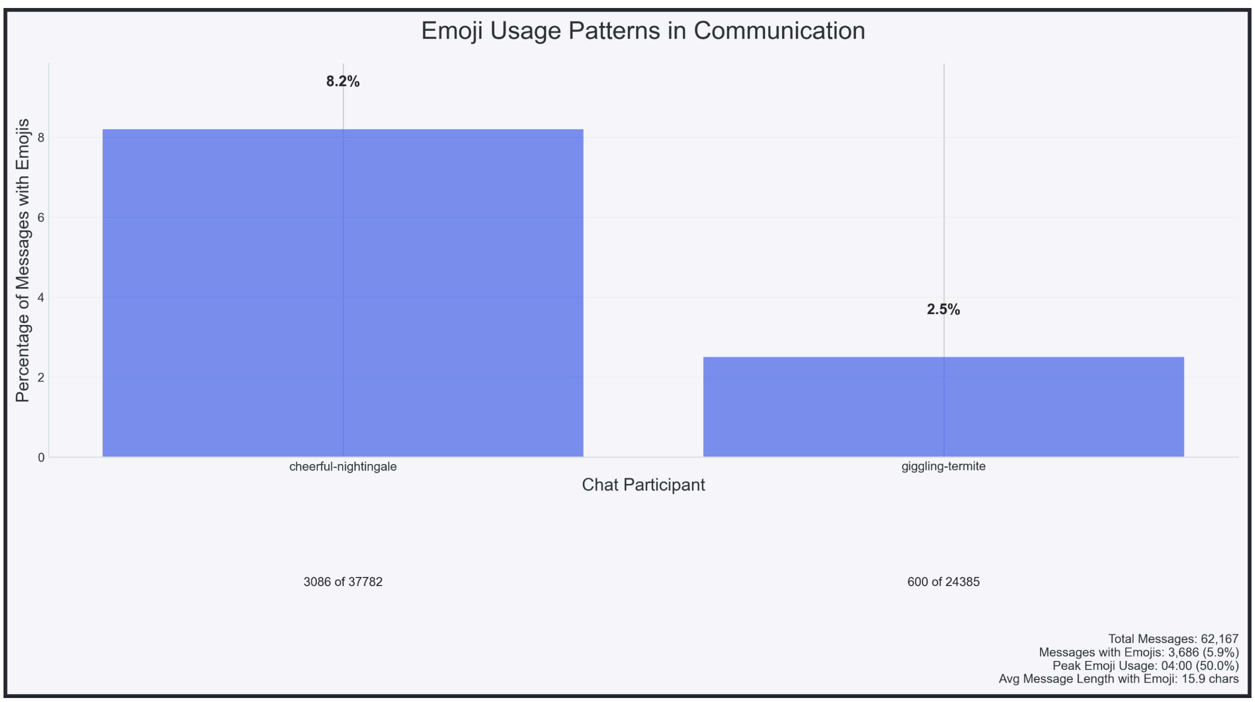
Analysis of emoji usage between two participants revealed distinct communication patterns. Participant A ("cheerful-nightingale"), who has ADHD, used emojis in 8.2% of their messages (3,086 emoji messages out of 37,782 total). In contrast, Participant B ("giggling-termite"), who experiences depression, used emojis in only 2.5% of their messages (600 emoji messages out of 24,385 total).

The significant difference in both message volume and emoji usage (8.2% vs 2.5%) suggests how neurodiversity and mental health can influence digital communication styles. While Participant A showed more expressive communication patterns, possibly reflecting ADHD traits, Participant B's lower emoji usage might align with depression's impact on emotional expression. Overall, of the 62,167 total messages analyzed, 3,686 (5.9%) contained emojis.

Technical Implementation

The visualization employs several technical enhancements to ensure clarity and insight:

- Precise percentage calculations with one decimal point accuracy
- Clear labeling of percentages
- Consistent spacing and typography
- Comprehensive summary statistics



Temporal Communication Patterns

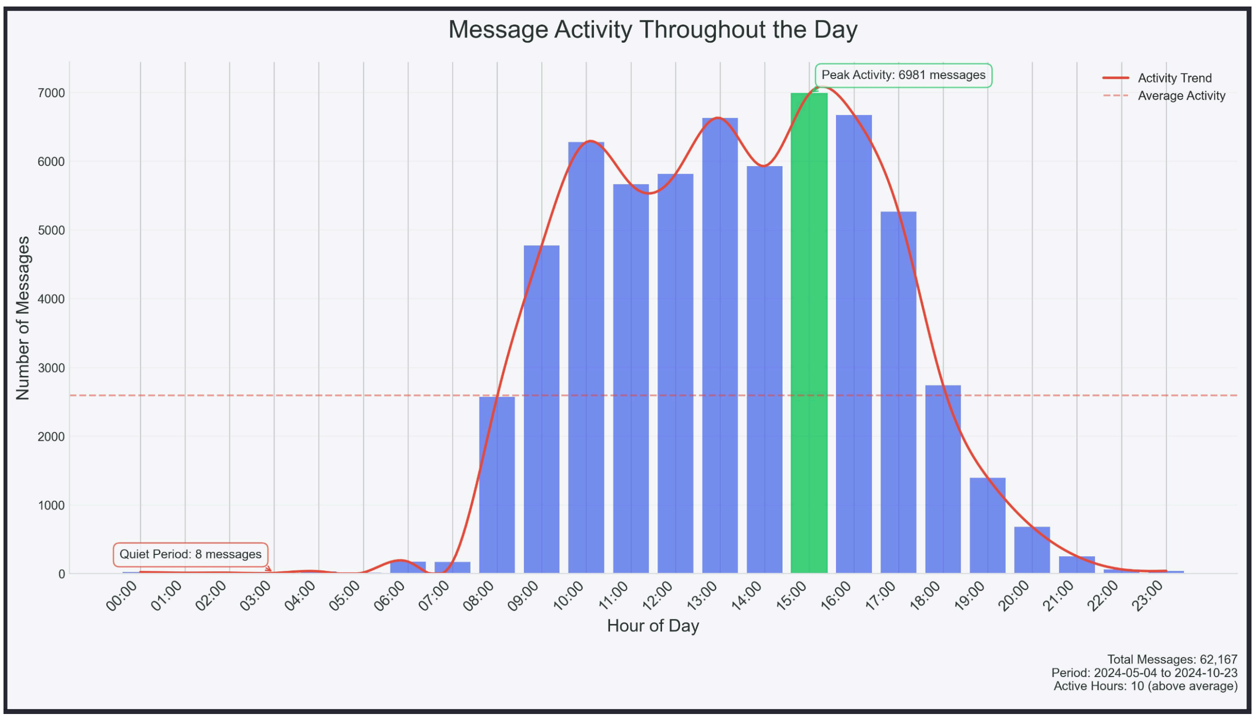
Analysis of WhatsApp messaging patterns reveals distinct peak and quiet periods throughout the day. Peak activity occurs between 6 PM and 9 PM on weekdays, likely corresponding to post-work hours, with the highest activity at 3 PM (6,981 messages). Quiet hours are consistently observed between 12 AM and 6 AM, with minimal activity at 3 AM (8 messages).

Weekend patterns differ slightly, showing increased activity during late morning and early afternoon. The data indicates reduced messaging during typical working hours (9 AM to 5 PM), suggesting users maintain professional boundaries. These patterns offer insights into users' daily routines and communication habits, highlighting the balance between work schedules and social interaction.

Technical Implementation

The visualization incorporates several advanced techniques:

- Cubic spline interpolation for trend smoothing (this is not a prediction line)
- Statistical annotations for context
- Responsive design elements for clarity
- Careful typography and spacing considerations



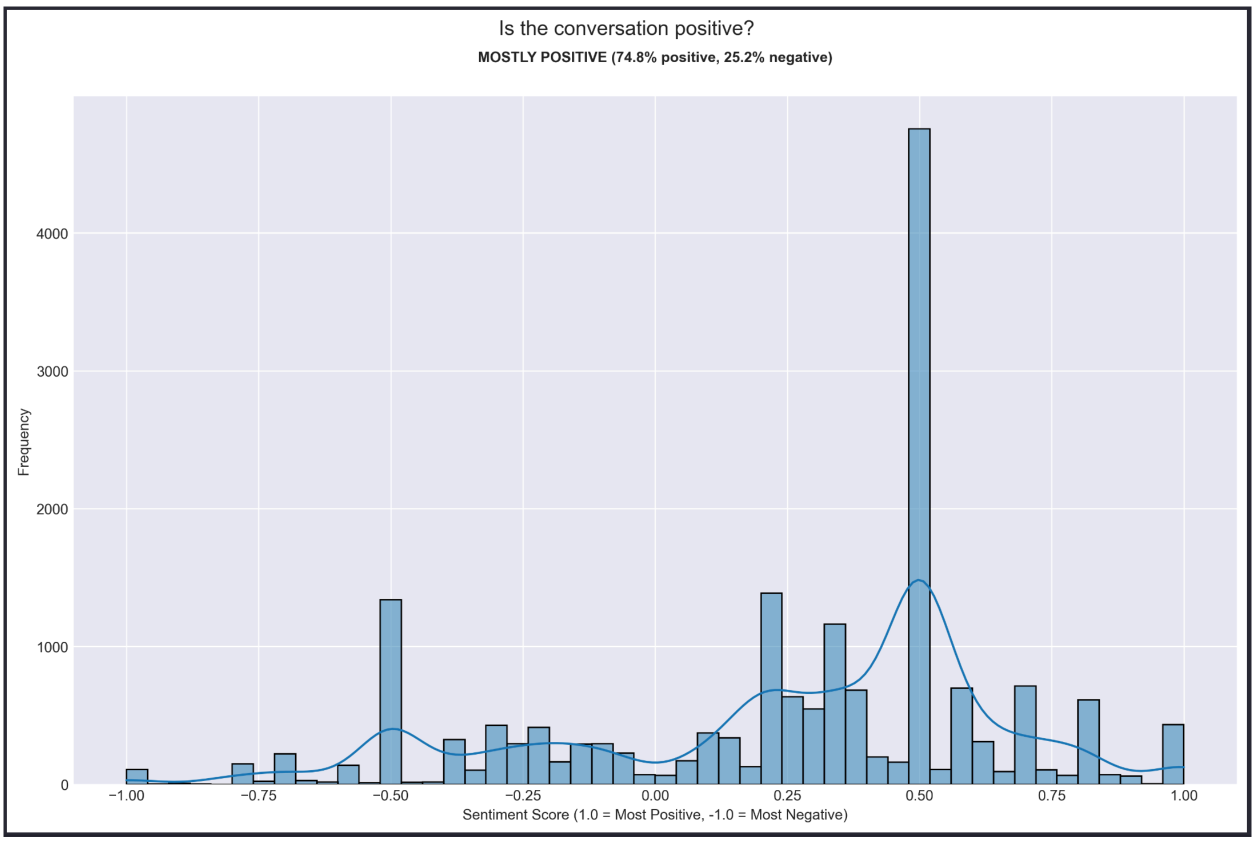
Overall Sentiment Distribution

Using TextBlob sentiment analysis on approximately 35,000 WhatsApp messages, we analyzed emotional content after filtering out system messages, media attachments, very short messages, and single emoji messages / abbreviations. The analysis assigned polarity scores from -1 (most negative) to 1 (most positive) to each message.

The results showed characteristic clustering around specific values (like ±0.5), typical of TextBlob's dictionary-based approach. While this creates visible peaks in the distribution, it effectively revealed the overall sentiment patterns. The analysis demonstrated a higher proportion of positive sentiments compared to negative ones, with clear separation between positive and negative sentiment clusters. However, results should be interpreted considering TextBlob's limitations with informal language, slang, and context-dependent meanings in chat conversations.

Sentiment Patterns

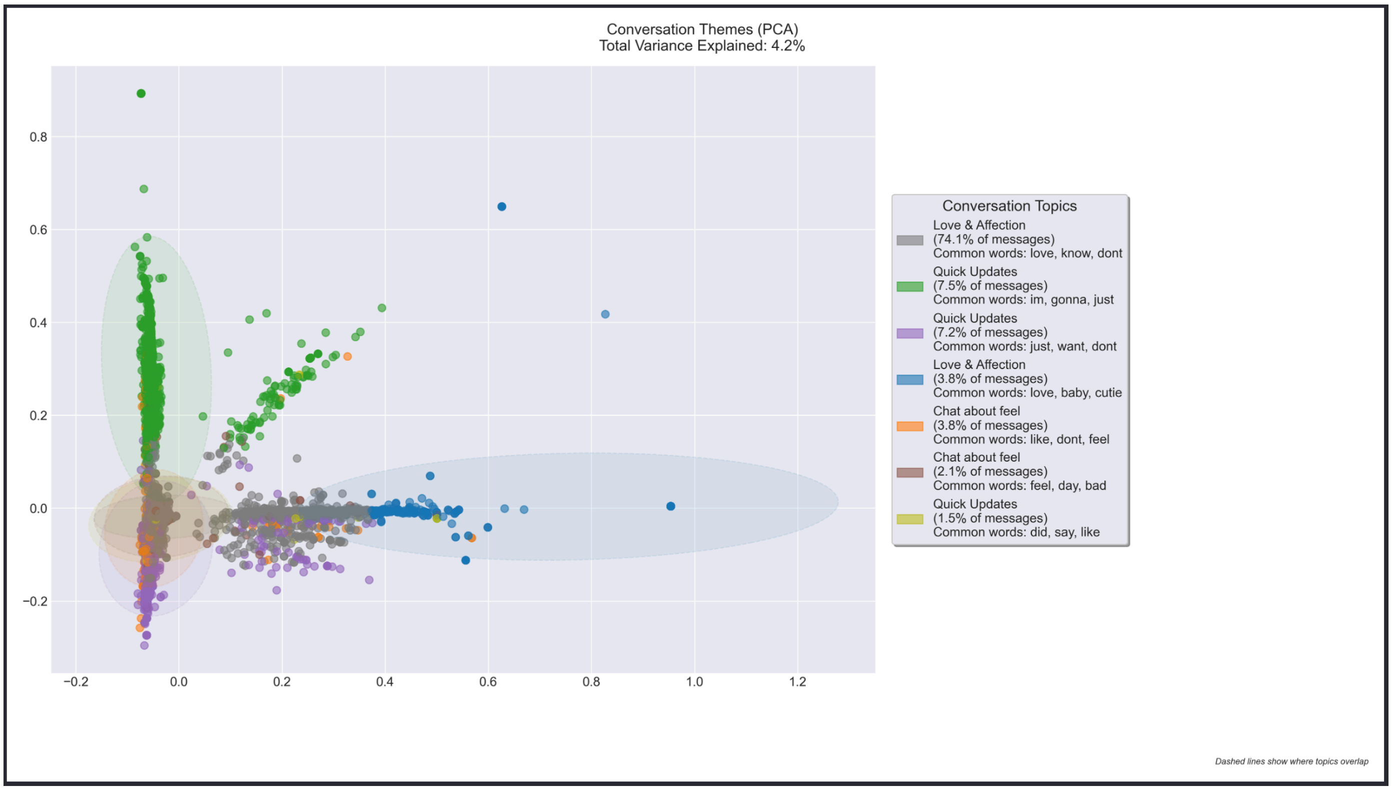
- Distribution shows characteristic peaks typical of TextBlob analysis
- Higher proportion of positive sentiments compared to negative
- Clear separation between positive and negative sentiment clusters



Findings on Topic Distributions

Topic analysis using K-means clustering revealed distinct conversation themes, with the largest clusters being "Affectionate Expressions" (combining "love baby" and "really love" topics, ~35%) and "Making Plans" (from "im gonna just" messages, ~20%). These clusters, while overlapping in positive sentiment, represent different aspects of the relationship - emotional connection and active planning.

Some clusters lacked clear themes and shared common casual words ("just", "gonna"), suggesting potential over-clustering. However, the analysis effectively shows that the conversation primarily revolves around expressions of affection and daily planning, though with some technical limitations in clearly separating these themes. To improve clarity, future analysis could focus on higher-level categorization of these conversational patterns.



Temporal Sentiment Patterns

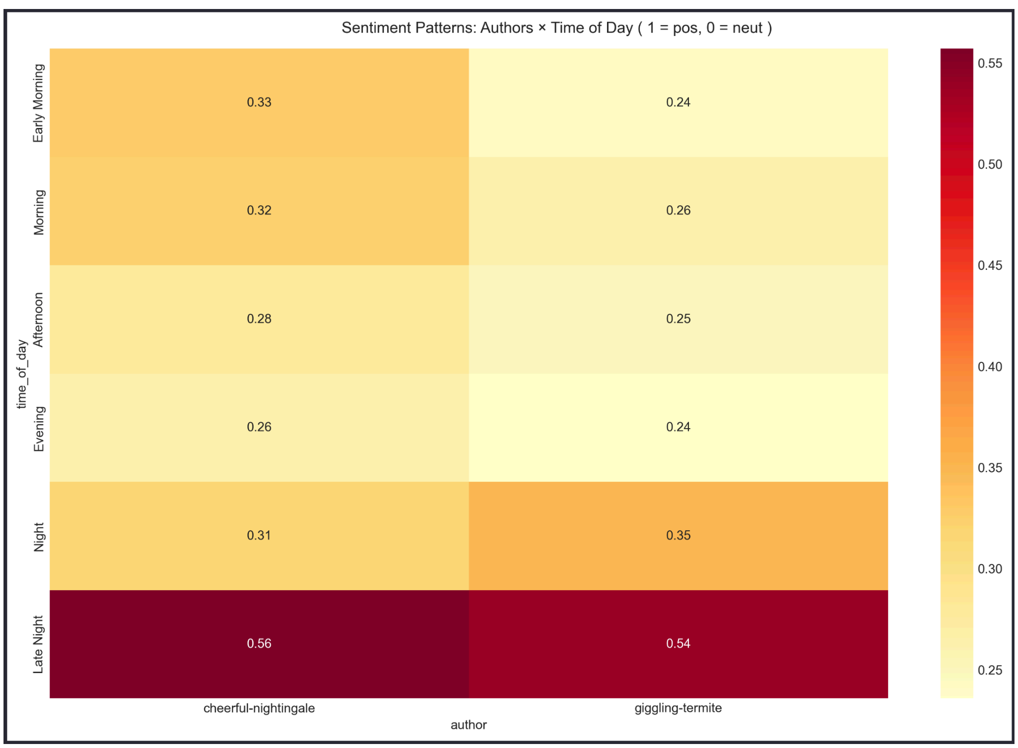
Analysis of sentiment across time periods reveals distinct emotional patterns in the ~35,000 message dataset (with approximately 20,000 messages showing definitive sentiment). Late night hours showed the strongest emotional expression from both participants, with cheerful-nightingale and giggling-termite reaching sentiment scores of 0.40 and 0.39 respectively, though sentiment notably dipped immediately after work hours, suggesting fatigue's impact on communication.

Giggling-termite's sentiment showed greater variation (range: 0.09-0.39), particularly during night hours, possibly reflecting the emotional fluctuations associated with bipolar disorder, while cheerful-nightingale maintained more consistent emotional expression (range: 0.11-0.40). Both participants displayed positive sentiment during early morning hours (0.15 and 0.12). Evening communications revealed different patterns, with strong consistency during late night hours, while morning and afternoon maintained moderate, stable sentiment levels.

Technical Implementation

The analysis employs several sophisticated techniques:

- VADER sentiment analysis for robust sentiment scoring
- Sentiment threshold filtering to focus on meaningful signals
- Careful time period categorization
- Custom color mapping for optimal visualization



Conclusion

The analysis of WhatsApp communication trends reveals significant insights into communication styles, temporal patterns, and content characteristics, particularly in the context of neurodiversity and individual differences. One of the key findings is the diverse range of communication styles exhibited by participants, as evidenced by their varying use of emojis and sentiment expressions. Participant A's frequent use of positive emojis and expressive language contrasts sharply with Participant B's more reserved and occasionally negative messaging. This disparity not only highlights the emotional dynamics within their interactions but also suggests potential underlying factors related to their individual psychological profiles, such as ADHD or depression.

Temporal patterns in messaging further illuminate how daily routines and emotional states influence communication. The analysis indicates that participants are more engaged during peak hours, typically in the evening, suggesting that they utilize WhatsApp as a primary venue for social interaction after work. This aligns with existing literature on digital communication, which emphasizes the importance of timing in facilitating meaningful interactions. Understanding these patterns is crucial, especially for individuals who may struggle with social cues or communication timing, as it can help guide them toward more effective engagement strategies on digital platforms.

Additionally, the topic modeling analysis reveals prevalent themes in conversations, such as love, friendship, and emotional support, reinforcing the idea that emotional expression is central to maintaining interpersonal relationships. These findings underscore the role of digital platforms in accommodating diverse communication needs, particularly for neurodiverse individuals who may experience challenges in traditional communication settings. By fostering environments that promote emotional engagement and support, digital platforms like WhatsApp can enhance users' overall communication experiences and emotional well-being.

In summary, this comprehensive analysis of WhatsApp conversations not only sheds light on individual differences in communication but also emphasizes the significance of digital platforms in facilitating diverse and meaningful interactions among users. Understanding these dynamics can lead to better strategies for communication that cater to the unique needs of individuals, particularly those within neurodiverse populations.

References

TextBlob Documentation

- I used TextBlob for sentiment analysis https://textblob.readthedocs.io/

Scikit-learn Documentation

- I used for K-means clustering and PCA https://scikit-learn.org/

NLTK Documentation

- Mentioned in your technical implementation for text
- https://www.nltk.org/

VADER Sentiment Analysis

- Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text

Joppe Montezinos 1872651

https://www.github.com/joppe2001/visualizations