Project Report

WhatsApp Chat Analyzer

Developed by

Saiful Islam Rupom

Introduction

With the exponential rise in instant messaging, WhatsApp has become one of the most widely used communication platforms globally. Billions of messages are exchanged daily, containing not only text but also media, links, and emojis. Despite its widespread use, WhatsApp does not provide users with a built-in way to analyze their chats for trends, usage statistics, or patterns of communication. The WhatsApp Chat Analyzer project addresses this gap by providing an interactive, web-based dashboard that enables users to upload their exported WhatsApp chat history and visualize insights in real-time.

Objective

The primary objective of this project is to design and develop a data analysis–based web application that allows WhatsApp users to:

- Analyze personal and group chats in detail.
- Visualize chat activity across different timeframes (daily, monthly, weekly).
- Identify active participants, common words, and emoji usage.
- Transform raw text exports into meaningful and interactive insights.

Methodology

The project follows a structured workflow, from data preprocessing to interactive visualization:

Step 1: Data Collection

- WhatsApp allows users to export chat histories (with or without media).
- For this project, .txt files exported without media were used.

Step 2: Data Preprocessing

- Implemented in preprocessor.py.
- Extracted essential features:
  - User names
  - Messages
  - Date and time
- Removed system messages like group_notification.
- Converted data into a clean Pandas DataFrame.

Step 3: Statistical Analysis

- Implemented in helper.py.
- Metrics computed:
  - Total number of messages
  - Word counts
  - Media shared
  - Links extracted

Step 4: Visualization

- Matplotlib & Seaborn for bar charts, timelines, and heatmaps.
- WordCloud for visual representation of most common words.
- Emoji package to extract and quantify emoji usage.

Step 5: Interactive Web App Development

- Built using Streamlit (app.py).
- Sidebar for:
  - Uploading WhatsApp .txt file
  - Selecting a user (or "Overall" for group analysis)
- Main dashboard for:
  - Quick stats (metrics)
  - Timelines
  - Activity maps
  - WordCloud & word frequency charts
  - Emoji analysis

Features

1. Quick Stats
- Total messages, word counts, media, and links displayed at a glance.

2. Timelines
- Monthly and daily activity visualized using line graphs.

3. Activity Maps
- Most busy day of the week.
- Most busy month.
- Heatmap showing weekly activity distribution.

4. Most Active Users
- Identifies top contributors in group chats.
- Presented as bar charts and data tables.

5. WordCloud & Most Common Words
- WordCloud visualization of frequent words.
- Horizontal bar chart showing most used words with counts.

6. Emoji Analysis
- Table showing emojis with usage counts.
- Pie chart displaying percentage of top 10 emojis.

7. Individual vs. Group-Level Analysis
- Ability to switch between analyzing one user or the overall group.
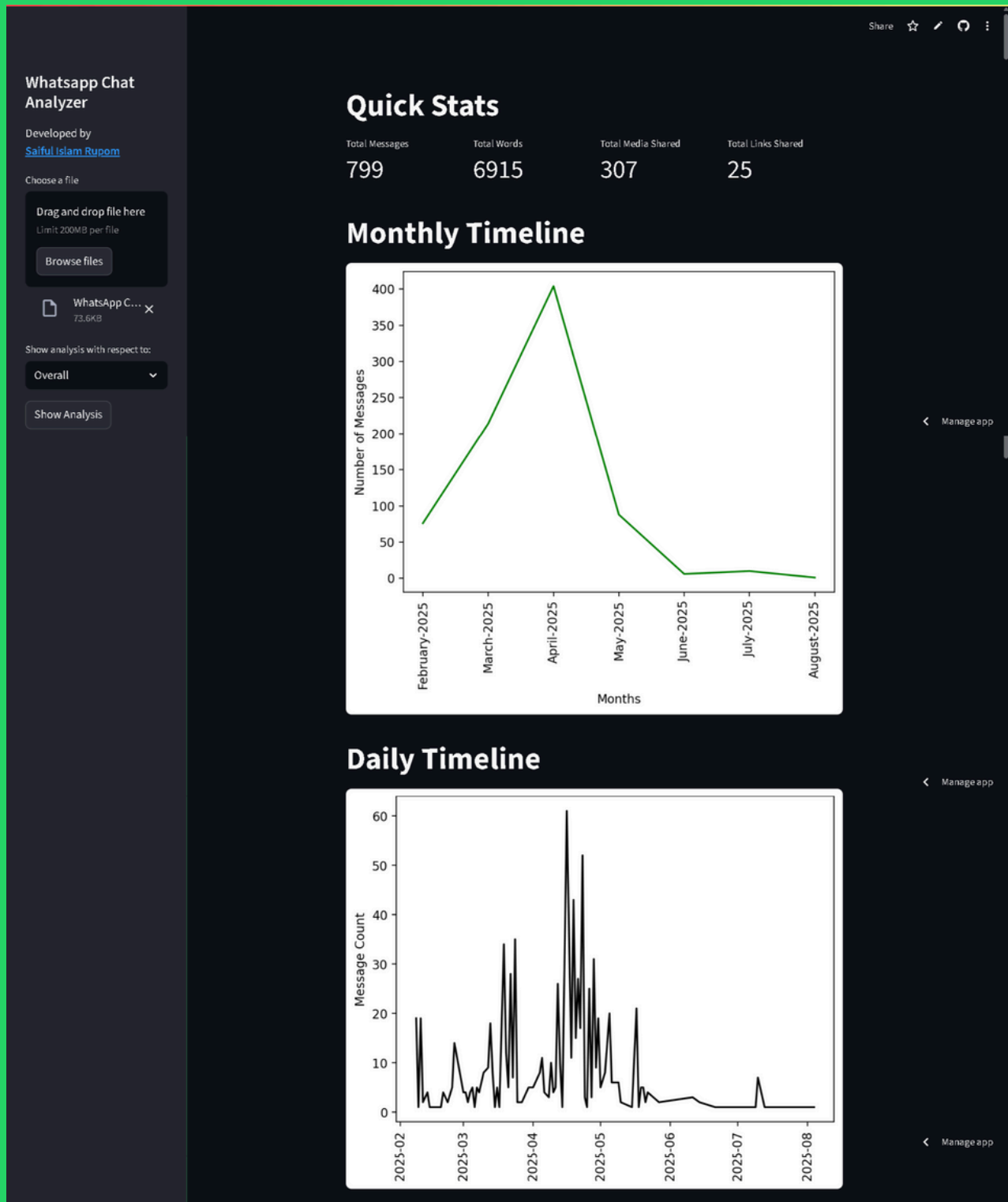
System Requirements

Hardware

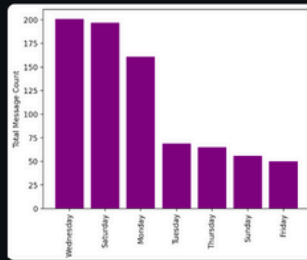- Any modern system with >4GB RAM.

Software

- Python 3.10+
- Required Libraries:
  - streamlit
  - pandas
  - matplotlib
  - seaborn
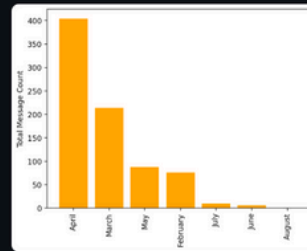 - wordcloud
  - emoji
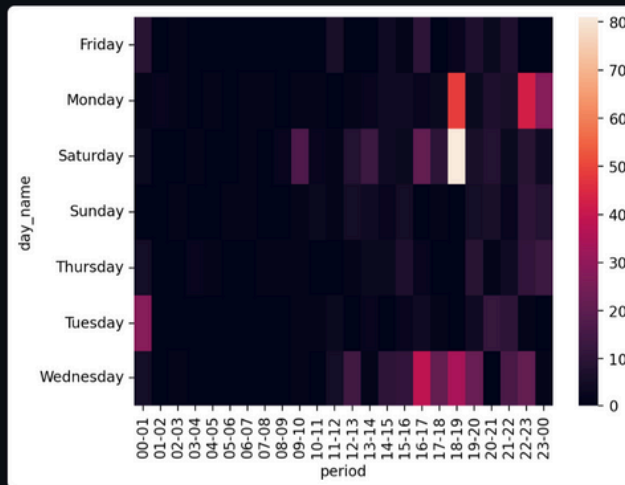 - urlextract

App Snapshots

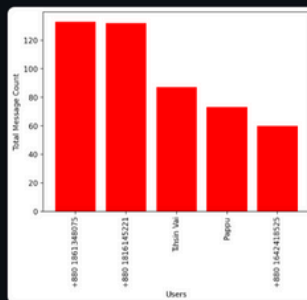# Activity Map

Most busy day

Most busy month

Weekly Activity Map

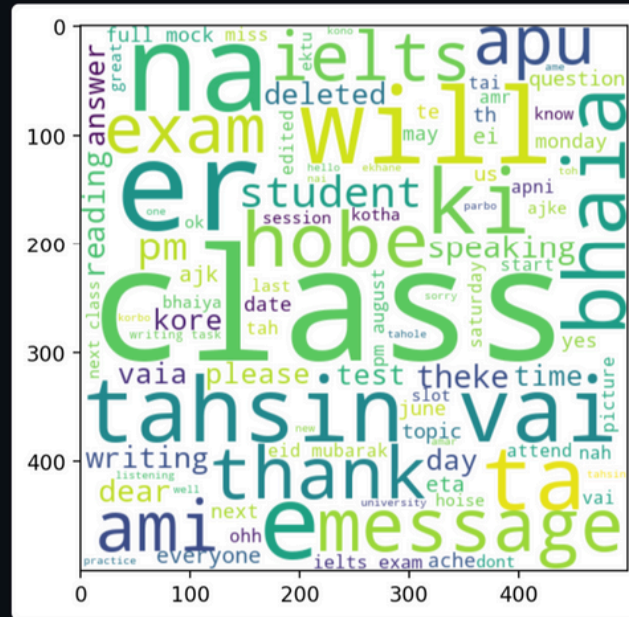# Most Active Users

	Person	Percentage
0	+880 1861-348075	16.65
1	+880 1816-145221	16.52
2	Tahsin Vai	10.89
3	Pappu	9.14
4	+880 1642-418525	7.51
5	+880 1629-880395	6.38
6	group_notification	5.88
7	+880 1768-748778	4.13
8	+880 1609-896897	3
9	+880 1897-209823	2.88

Wordcloud

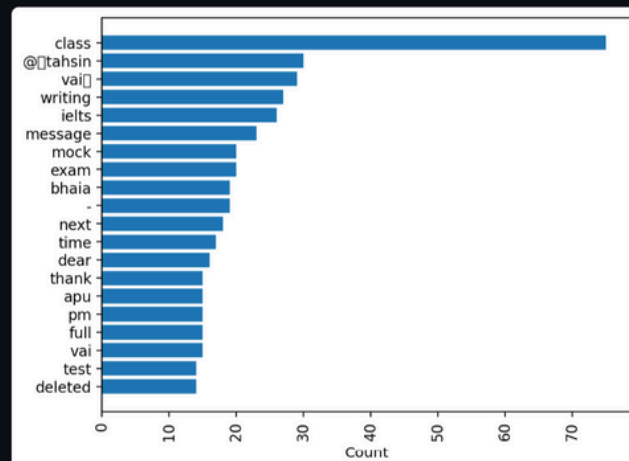
Most common words

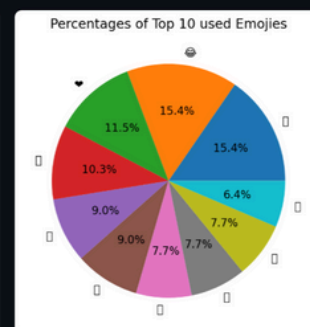## Emoji Analysis

	Emoji	Count
0	😂	12
1	🤣	12
2	❤️	9
3	🌙	8
4	📱	7
5	🔥	7
6	😊	6
7	🗓️	6
8	🎯	6
9	👀	5

Percentages of Top 10 used Emojies

15.4% · 15.4% · 11.5% · 10.3% · 9.0% · 9.0% · 7.7% · 7.7% · 7.7% · 6.4%

Thank you for exploring your chats with this tool.

Click here to try the App

Future Enhancements

- **Sentiment Analysis:** Categorizing messages as positive, negative, or neutral.
- **Topic Modeling:** Identifying conversation themes automatically.
- **Network Graphs:** Visualizing interaction between participants.
- **Report Generation:** Export insights as PDF or CSV.

Conclusion

The WhatsApp Chat Analyzer successfully demonstrates how raw chat data can be transformed into meaningful insights using data analysis and visualization techniques. By combining the power of Python libraries and Streamlit, the project provides an accessible, user-friendly tool that enhances digital communication awareness.

This system has potential applications in:

- Personal productivity tracking
- Group communication analysis
- Academic studies of online communication patterns