WHATSAPP COVERSATION EVALUATOR

ABSTRACT

WhatsApp (officially as WhatsApp Messenger) is a voice-over-IP(VoIP) and an instant messaging (IM) application where it gives us the option to ship textual content, video messages and voice messages. WhatsApp chat evaluation ambitions to extract treasured statistics, patterns, and insights from the messages exchanged on the platform. it could serve distinct capabilities, which includes knowledge consumer conduct, sentiment analysis, monitoring agency dynamics, and extracting applicable facts for preference-making technique: The paper possibly outlines the method hired for amassing and analysing WhatsApp chat facts. this will include guide coding of conversations, automatic textual content analysis techniques, or a mixture of every Implications: The findings of the analysis may have implications for know-how digital communication patterns, social relationships mediated with the aid of generation, and the characteristic of language in shaping on line interactions

WhatsApp is the most used communication method and it is also useful. There are many types of corporate and personal communication. So, there may be some truth unseen in it. This project engages these discussions and provides an in-depth analysis of these documents. No matter what the topic is, discussion ensures a good and accurate evaluation of the paper. The main benefit of this project is seaborn, emoji, matplotlib, pandas etc. is to use libraries to create statistical methods and graphs in a green way.

key phrases: Communication done in WhatsApp, Pandas, Evaluation, Natural Language Processing (NLP), Python, Matplotlib, Front-end Streamlit, Emoji.

I. INTRODUCTION

In the given document, we were able to create a WhatsApp speech analyser. WhatsApp chats give various modes of communication among group members and individual participants. Sent text messages can be used in many NLP and machine learning models. These creative ideas provide the necessary training. The app can check that kind of information from WhatsApp chat transaction. The main benefit of this structure is the use of core libraries from the python for instance seaborn, streamlit, pandas, NumPy, matplotlib. They are often optimized to create documents and graphics. Pre-existing data does a role that is very efficient and important in the sub-field of artificial intelligence that is machine learning. Us here only focus on one application that is WhatsApp, which is among Facebook's Big data generator, as they require a lot of information to build an ideal model. WhatsApp assures that 50 billion messages are transmitted every day. More than 500 minutes per week is the average time that a user spends on WhatsApp.

II. Problem Statement

WhatsApp communication evaluator is an analytical tool for conversation that takes place in WhatsApp. Running the data export dialog will help you create unique metrics. For example, most of the time with other groups of people participating in the conversation. We recommend using the data management system to better understand WhatsApp chats on your phone.

Literature Survey

2.1 Literature overview on WhatsApp Chat evaluation:

We conducted an analysis on WhatsApp Messenger Usage and Stress Research [1] and conducted a comprehensive research and analysis. This research covers the impact of WhatsApp on university students (young people). Many recommend using WhatsApp on other websites. They exchange images, sound

and video. The survey also proves that WhatsApp is the most used smartphone app than any other app. This research was conducted on the positive and negative data of WhatsApp usage. As we know from this survey, WhatsApp is the most used application among young people and generations, so our business can give them insight and provide unknown information during the conversation.

2.2 Module evaluation data:

A. Matplotlib: A famous Python software package used to visualize integrated circuits is Matplotlib. It is a cross-platform library for creating merged images from merged sequences. Allows integration in statics, animations and interactive visualizations in Python

B. Streamlit: A free and open-source Python framework is Streamlit. [2] We can use Streamlit to accelerate network services that integrate get-built-integrated and actual technologies. Streamlit integrates seamlessly with other popular Python applications such as NumPy, Pandas, Matplotlib, Seaborn, etc. Streamlit provides the fastest way to create and deliver integrated packages.

C. Word Cloud: A word cloud is a library that creates visual information used to represent the most common and repeatedly used expressions in text. Frequently used and important phrases are shown in capital, bold letters.

D. Seaborn: Seaborn is a visual library. Mileage is used to create integrated statistical graphs. Visualization is an important part of seaborn. Seaborn has an integrated approach to research and better understanding. Seaborn is tightly integrated into Python's data structure.

E. Pandas: Pandas is a Python library which is open source. Pandas is used to convert str-built structures into real bodies. The actual frame is a built-in representation of a two-dimensional table consisting of rows and columns. We can draw using the built-in Pandas library. Pandas -library has many evaluations,

integration, information discovery and optimization functions because pandas want to achieve all the success. Tailored Built-in assessment tools

2.3 Python Literature evaluate:

Python is a general-purpose language. It has an easy-to-learn syntax. Python is a green and useful language that can be used for scientific research in the discipline of computing and real-time analysis, allowing programmers to improve their skills. It is designed for mileage analysis and has many features for quick visibility. Python is loose and open; Giving you the tools you need is the key to the best research. Unlike MATLAB or LabView, Python can be used in any programming project. Scientists work with ancient and complex languages, so they need Python tools to help them analyse them easily. Python is dedicated to improving and blessing professors and college students (Gergely, I., 2014). Tony, J. (2004) Conducting an experiment using Python as a first language. According to experienced researchers, solving complex learning problems using C++ takes several hours, while students using Python take less than an hour. Python is advanced, flexible and powerful and can be used for many types of business. Python has good support for function types and has a large and complete library. (Srinath, k.R., 2017) This research was initiated and the Python interpreter was developed for Windows, Linux, UNIX etc. It was found to be available for various operating systems such as.

2.4 Evaluation of Internet Design Literature:

The number of Internet users has reached tens of millions and is expected to increase in the next few years. Websites are an important medium for recording, publishing and broadcasting. The purpose of this article is to review previous

research on web development. This document therefore warns against a fixed pattern or change in service technology and infrastructure (especially network connections). The acceptance and success of websites and e-commerce depends on website design. The purpose of this article is to identify and understand user beliefs and behaviours towards utilizing e-commerce sites. According to research (Lee and Kozar, 2012), there is currently no agreement on how to properly operate a website and verify its usability. Nielsen associate's usability with learning, performance, memory, error, and enjoyment (Nielsen, 2012). We don't currently have a guide on how to create a website to attract customers. Hyperlinking is the process of creating web pages that embed content and link to the Internet. Documents written in HTML or XML. CSS describes how elements in web design are displayed on the screen, in text, in speech, or in other media.

Methodology

3.1 Data Analysis

This is a method of cleaning, reworking, analyzing and showing facts to learn some useful analysis

and then share some decisions. The way it is decomposed divides all the elements into separate additives for one to examine. Fact testing is a way to take facts and turn them into useful information that consumers can choose from. This project provides basic information about WhatsApp Chat. Here is the review:

* Find all messages, similar messages, all news and links shared in WhatsApp chats
* Find the most active people in your relationship.
* Find the busiest and least busy days of the month.

3.2 Preparation process

Preliminary information is the starting point of this project, which is about understanding how to use and use the different

built-in Python modules. These different models enable users to better understand and represent their rights. Statistics such as all comments, all news, links, shared images, game maps

Weekly and monthly maps, monthly program, daily program, most users, most used words,

used Emojis.

3.3 Functionality

Steps to chat exportation:

Open WhatsApp group chat -> Click on Menu -> Click Connect -> Select Export chat -> Select without media. Perform WhatsApp chat analysis.

1)First, open the WhatsApp Chat Analyzer website.

2)Choose a day pattern.

3)Load exported chat history.

4) Use training models for data analysis

5) Data preprocessing is done by the learning version.

6)Choose a standard or free character test

7)Analytical models include peak statistics, cloud charts, interest rates, monthly ranges, daily ranges, score reports

3.4 System Modules

(a) Installing and importing dependencies: In this step install and import matplotlib, Streamlit, seaborn, collections, pandas, emoji, Wordcloud, URLextract and regular expression. (b)This step completes the preprocessing of the data. The information here is formatted and separated by date, time, username and reference information.

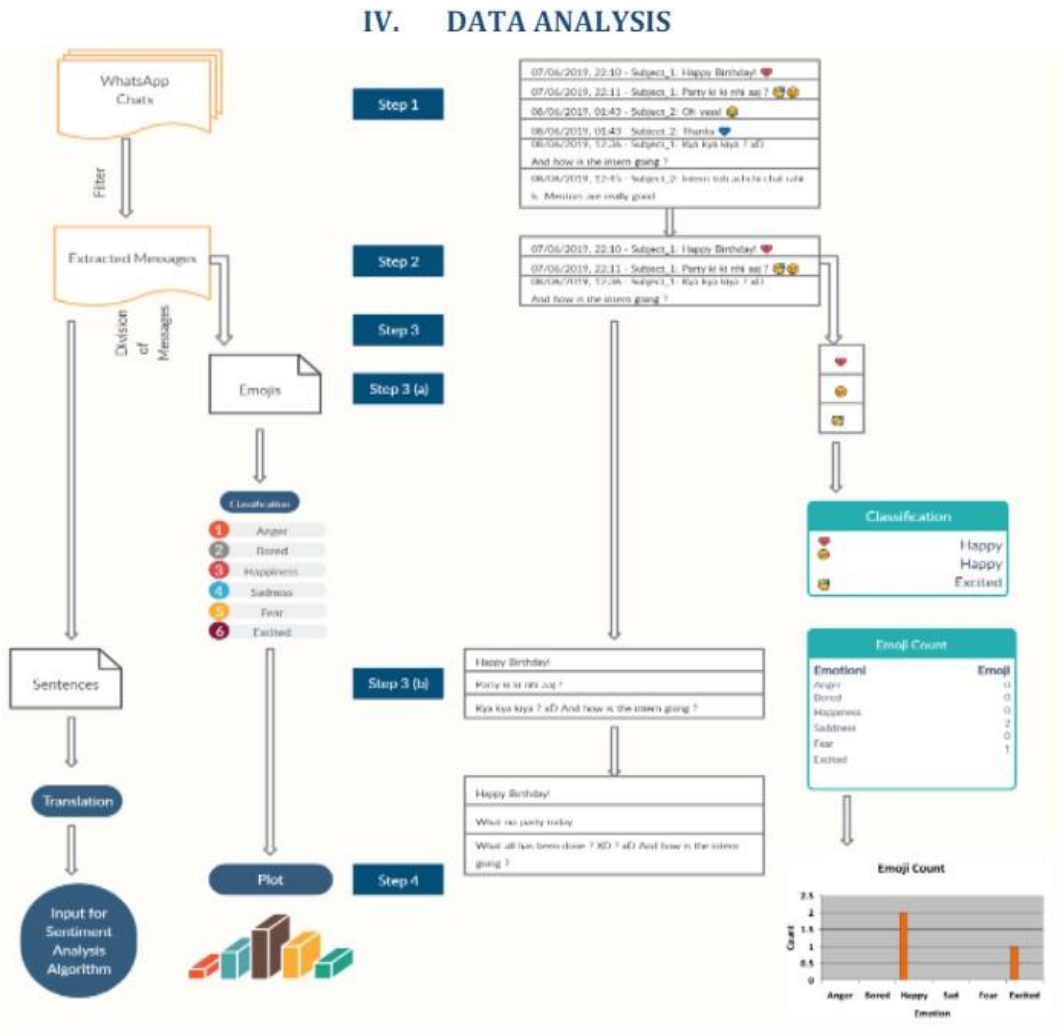
(c) Open individual or group chat->Click Options→Next→Export Chat->Select Export

No message->File set. Load the chat archive and click Analyze

(d) Show the chat model and analyze the data: The collected data is read and processed here for training the machine learning model. The sample is then measured and serialized. The review was carried out as follows:

1. Key statistics: total comments, total comments, including posts and links.
2. Monthly Meeting Frequency: Number of sessions per month.
3.Daily Schedule: Daily Discussion
4. Event Map: Show the busiest and least busy days of the same month
5. Weekly game plan. Word Cloud: Commonly used words. Biggest customer: Mostly employees. Emoji Analysis: Commonly used and used emojis.

4.Data Analysis



Here the raw data is converted into bits and stored in a variable ,then the variable is encrypted and stored into a table. Then manipulation is done on the data that is stored in the table according to the needs. Top statistics ,emojis used ,time line ,heat map where derived using various libraries inbuilt with the python programming language

5.Validation

Software validation is similar to a survey to determine the quality of the product being tested. The verification process involves running a program or application for a certain period of time, looking for software errors or vulnerabilities in it. The sample test cases of this project work are given in the table below

Sl. No	Test Case	Description	Expected Result	Test Result
1	Word Cloud	Frequent and most come used words	Most used word are produced in a word cloud	PASS
2	Activity Map	Least busy and most busy day is projected	Most busy and least busy day bar graph	PASS
3	Monthly Timeline	Chat is every month and its repetition	Graphs the timeline of months	PASS
4	Top Statistics	These are about all messages, all words, all news and links	Statistics of all news in the group, number of comments, number of links and news sharing numbers	PASS
5	Heat Map	A heat map visually represents data using colors to indicate intensity or magnitude	Produce heatmap output using a data visualization library consisting of Matplotlib or Seaborn in Python	PASS

6)RESULTS AND DISCUSSION

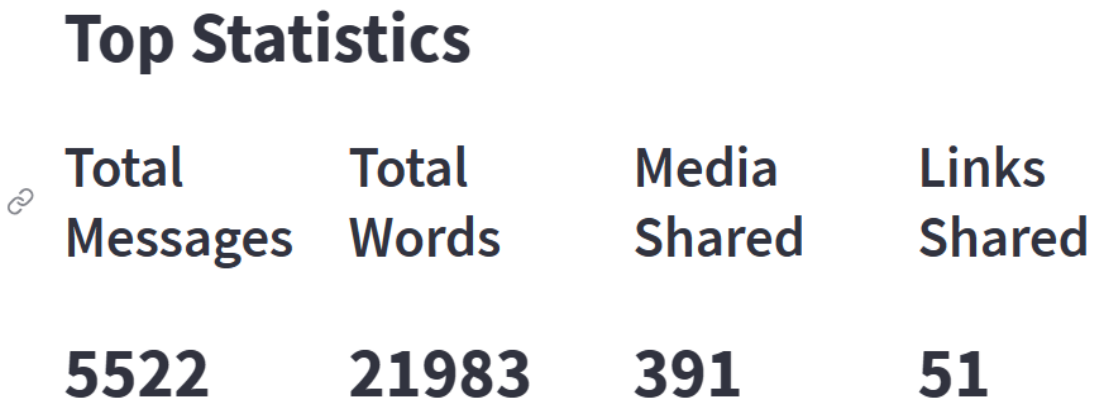


Fig 1. Top Statistics

It gives the overall statistics such as total words, total messages, media shared and link shared in each chat conversation

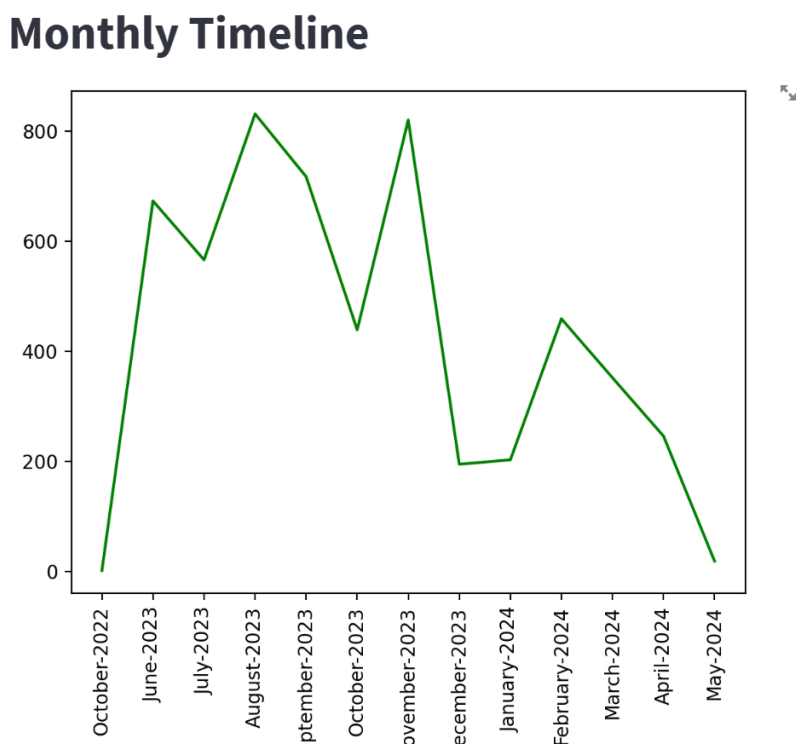


Fig 2 .Monthly Timeline

It gives the month where the conversation were most active and the month were the conversation were least active using visualisation

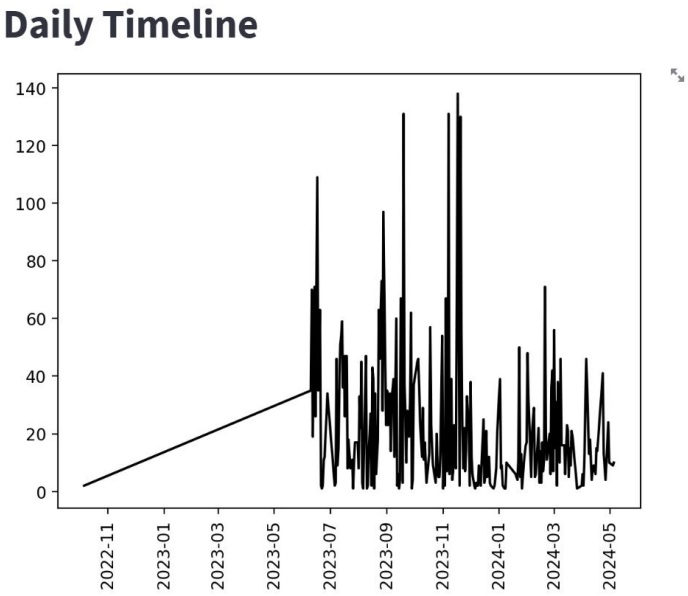


Fig 3

It gives us the statistics of when the most conversation in a day and the least busy day in the communication

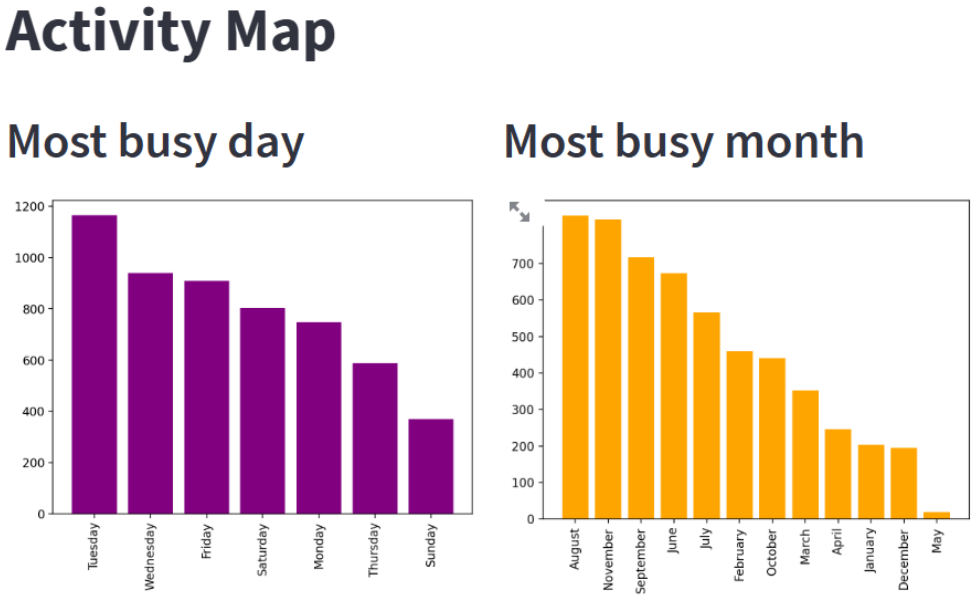


Fig 4

It gives us the busiest month and most busy day in the chat using bar plotting using matplotlib

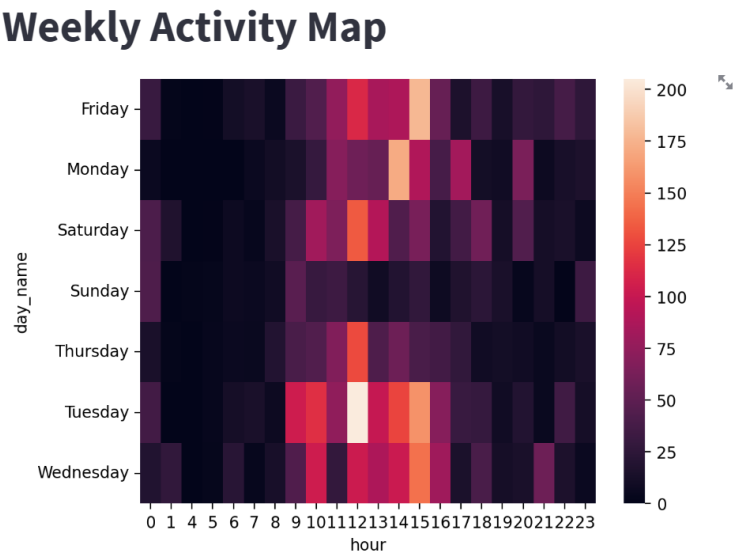
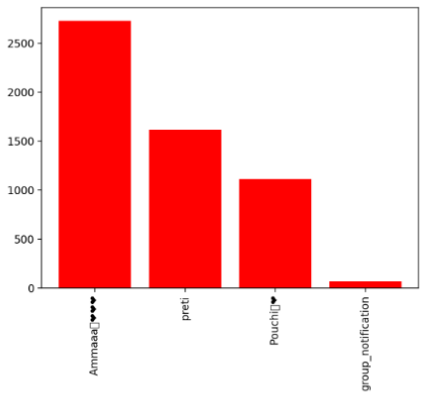


Fig 5

The weekly activity heatmap shows the distribution of activity by day of the week using color intensity.

Most Busy Users



	name	percent
0	Ammaaa❤❤❤	49.38
1	preti	29.23
2	Pouchi❤	20.14
3	group_notification	1.25

Fig 6

Here it plots in bar graph representing the member who contributes in communication a lot to least member who rarely participates in the group conversation

Wordcloud

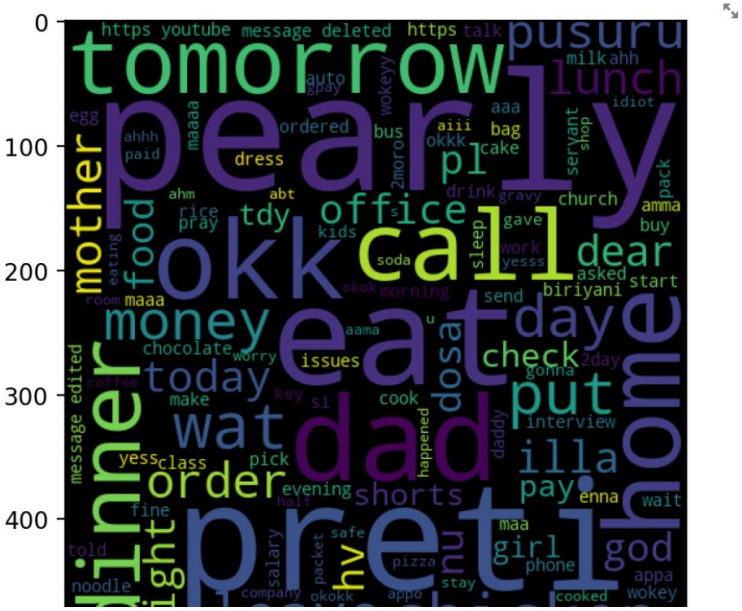


Fig 7

Word cloud give the accumulative and most common words used in the chatting increased font means most usage

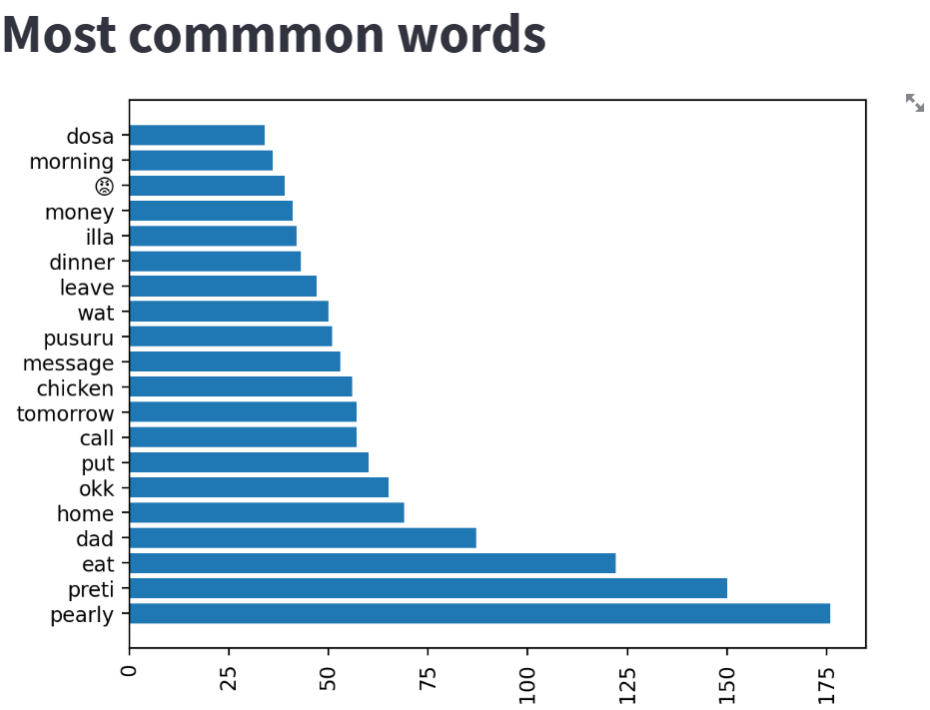


Fig 8

It analyses the entire chat and gives the most used word in each group chat. It uses matplotlib to do analysis

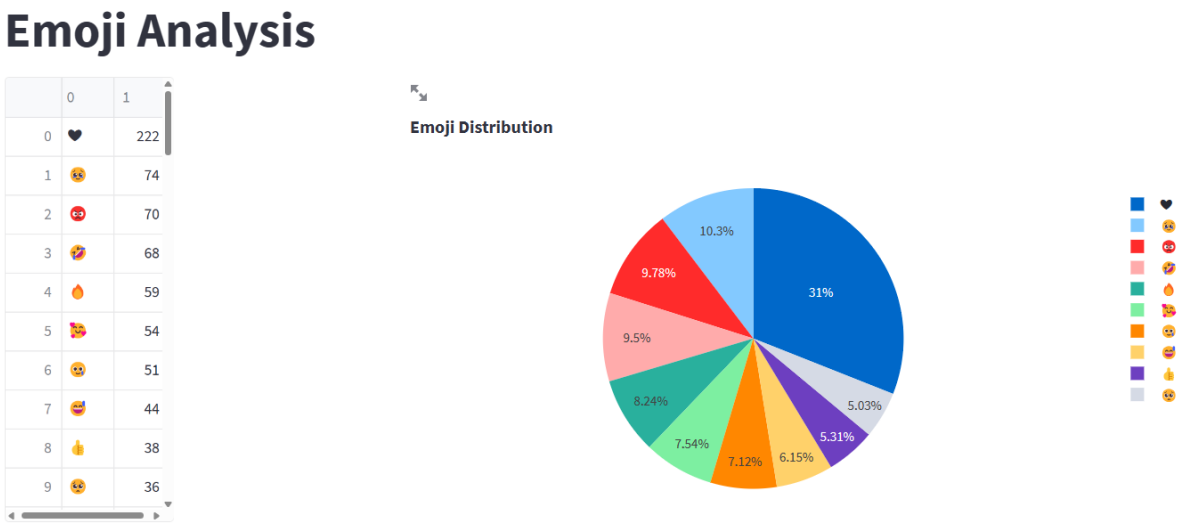


Fig 9

It gives us the emojis that the group have been using from most used emojis to least most emojis

7)Conclusion

We are going to conclude saying that this WhatsApp Chat Analyzer will tell us about the user behaviour, how much messages they have contributed in the group ,their most active day and month, the frequently used word and their favourite emoji to use. This gives us the data that we can use to analyse great insight in our future needs. The system implements Python and uses Python libraries including Streamlit, Emoji, NumPy, Pandas, Re, Matplotlib, URLextract, collections, and Seaborn. We finally achieved what we wanted. Our project will be important for organizations in the future. Then he finds out who is active in the group and who is active. They can decide accordingly.

8)References

[1] Ravishankara K, Dhanush, Vaisakh, Srajan I S, “International Journal of Engineering Research &
Technology (IJERT)”, ISSN: 2278-0181, Vol. 9 Issue 05, May-2020

[2] https://www.analyticsvidhya.com/blog/2021/06/build-web-app-instantly-for-machine-learningusing-streamlit/

[3] Dr. D. Lakshminarayanan, S. Prabhakaran, “Dogo Rangsang Research Journal”, UGC Care Group I
Journal, Vol-10 Issue-07 No. 12 July 2020

[4] https://www.interaction-design.org/literature/topics/web-design

[5] Marada Pallavi, Meesala Nirmala, Modugaparapu Sravani, Mohammad Shameem. WhatsApp Chat Analysis. International Research Journal of Modernization
in Engineering Technology and Science. Volume: 04/Issue:05/May-2022

[6] Shaikh Mohd Saqib. Whatsapp Chat Analyzer. International Research Journal of Modernization in Engineering Technology and Science. Volume: 04/Issue:05/May-2022

[7] K, Ravishankara & Dhanush, & Vaisakh, & S, Srajan. (2020). WhatsApp Chat Analyzer. International Journal of Engineering Research and. V9.10.17577/IJERTV9IS050676.
[8] D.Radha, R. Jayaparvathy, D. Yamini, “Analysis on Social Media Addiction using Data Mining Technique”, International Journal of Computer Applications (0975 – 8887).
[9] https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664
[10] https://www.analyticsvidhya.com/blog/2021/06/build-web-app-instantly-for-machine-learning-using-streamlit/
[11] Meng Cai, “PubMed Central”, PMCID: PMC7944036, PMID: 33732917
[12] E. Larson, "[Research Paper] Automatic Checking of Regular Expressions," 2018 IEEE 18th International Working Conference on Source Code Analysis and
Manipulation (SCAM), 2018, pp. 225-234, doi: 10.1109/SCAM.2018.00034.