

# Sentiment Analysis of Cultural Heritage Texts:

Anne Frank Diary of Young Girl

By: Smita Sanjay Pable

Advisor: Christina Moir

## Introduction

Synthesis of historical texts by NLP, as shown in Anne Frank's diary, brings out amazing information about the history of mankind. Anne's diary, a witness to her emotional journey, revealed her inner conflict. NLP algorithms along with the NRC Emotion Lexicon help to elucidate Anne's inner emotions and Streamlit shows us her courage and pivotal moments. Apart from individual emotions, this research adduces the patterns of psychology that shape the course of historical events. Anne's transcendent strength is what is left that will contribute to the improvement of cultural traditions and development of cultural identity.

## Preprocessing and Data Cleaning

The text preprocessing stage focuses on eliminating unnecessary characters and symbols, removing non-linguistic content such as dates and numbers, and filtering out stop words to facilitate meaningful analysis. Afterwards, the text is tokenized into discrete words for further analysis.

- Results:

SATURDAY, JUNE 20, 1942

Dearest Kitty! Let me get started right away: it's nice and quiet now. Father and Mother are out and Margot has gone to play Ping-Pong with some other young people at her friend Trees's. I've been playing a lot of Ping-Pong myself lately. So much that five of us girls have formed a club. It's called "The Little Dipper Minus Two." A really silly name, but it's based on a mistake. We wanted to give our club a special name: and because there were five of us, we came up with the idea of the Little Dipper. We thought it consisted of five stars, but we turned out to be wrong. It has seven, like the Big Dipper, which explains the "Minus Two." Ilse Wagner has a Ping-Pong set, and the Wagners let us play in their big dining room whenever we want. Since we five Ping-Pong players like ice cream, especially in the summer, and since you get hot playing Ping-Pong, our games usually end with a visit to the nearest ice-cream parlor that allows Jews: either Oasis or Delphi. We've long since stopped hunting around for our purses or money -- most of the time it's so busy in Oasis that we manage to find a few generous young men of our acquaintance or an admirer to offer us more ice cream than we could eat in a week.

Fig 1: Before Data Preprocessing and Cleaning

saturday june 20 1942  
dearest kitty let get started right away itsnice quiet father mother margot gone play pingpong young people friend treess ive playing alot pingpong lately much five us girls formed aclub itscalled little dipper minus two areally silly name itsbased amistake wanted give club aspecial name five us came idea little dipper thought consisted five stars turned wrong seven like big dipper explains minus two ilse wagner apingpong set wagners let us play big dining room whenever want since five pingpong players like ice cream especially summer since get hot playing pingpong games usually end avisit nearest icecream parlor allows jews either oasis delphi weve long since stopped hunting around purses money time itsso busy oasis manage find afew generous young men acquaintance admirer offer us ice cream could eat aweek youre probably alittle surprised hear talking admirers atender age unfortunately case may vice seems rampant school soon aboy asks bicycle home get talking nine times ten ican sure hell become enamored spot wontlet sight asecond ardor eventually cools especially since ignore passionate glances pedal blithely way gets bad start rambling asking fatherspermission iswerve slightly bike schoolbag falls young man feels obliged get bike hand bag time ive switched conversation another topic innocent types course blow kisses try take hold arm theyre definitely knocking wrong door iget bike either refuse make use company act iminsulted tell uncertain terms go home without weve laid basis friendship tomorrow anne

Fig 2: After Data Preprocessing and Cleaning

## Flow Chart of Sentiment Analysis

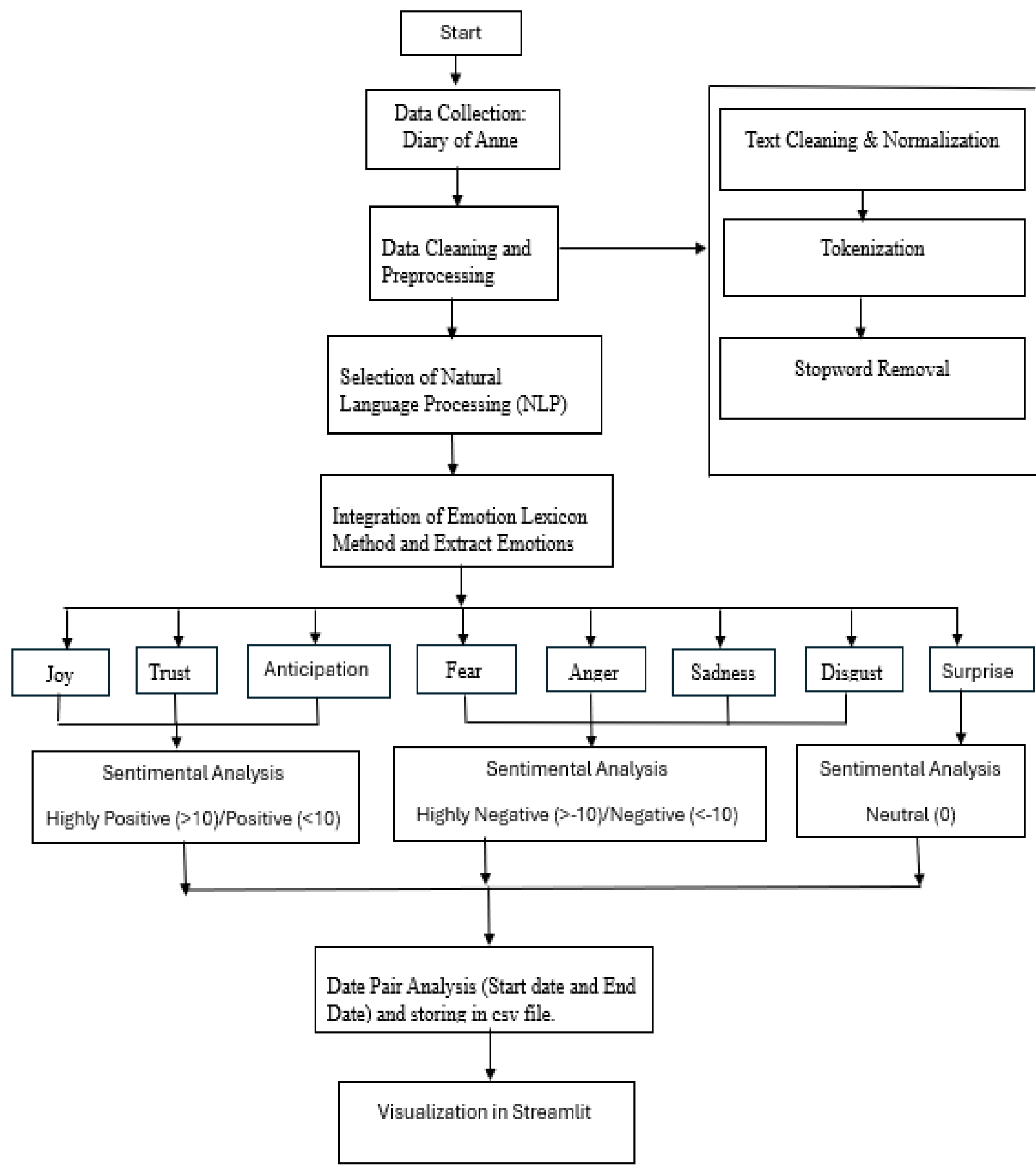


Fig 3: Flow Chart of Sentimental Analysis of Historical text

## Methodology for Sentiment Analysis and Emotion Extraction

- Data Collection: Obtaining Anne's Frank's diary entries.
- Text Preprocessing: Processing and formatting of text data for the analysis.
- NRC Emotion Lexicon Loading: Loaded the NRC Emotion Lexicon in which the words are marked with the emotions and sentiments rating.
- Sentiment Score Calculation and Labeling: The difference between the positive and negative emotions can be measured. Thus, this helps in calculating the sentiment scores and you can place them in clearly defined categories.
- Date Pair Analysis: Pick the date-entries in the diary text to know the sentiments and emotions trends over time.
- CSV Data Export: Save extracted emotions and sentiment scores along with the date entries in a CSV file for further analysis.
- Visualizing the extracted data in Streamlit.



## Visualization

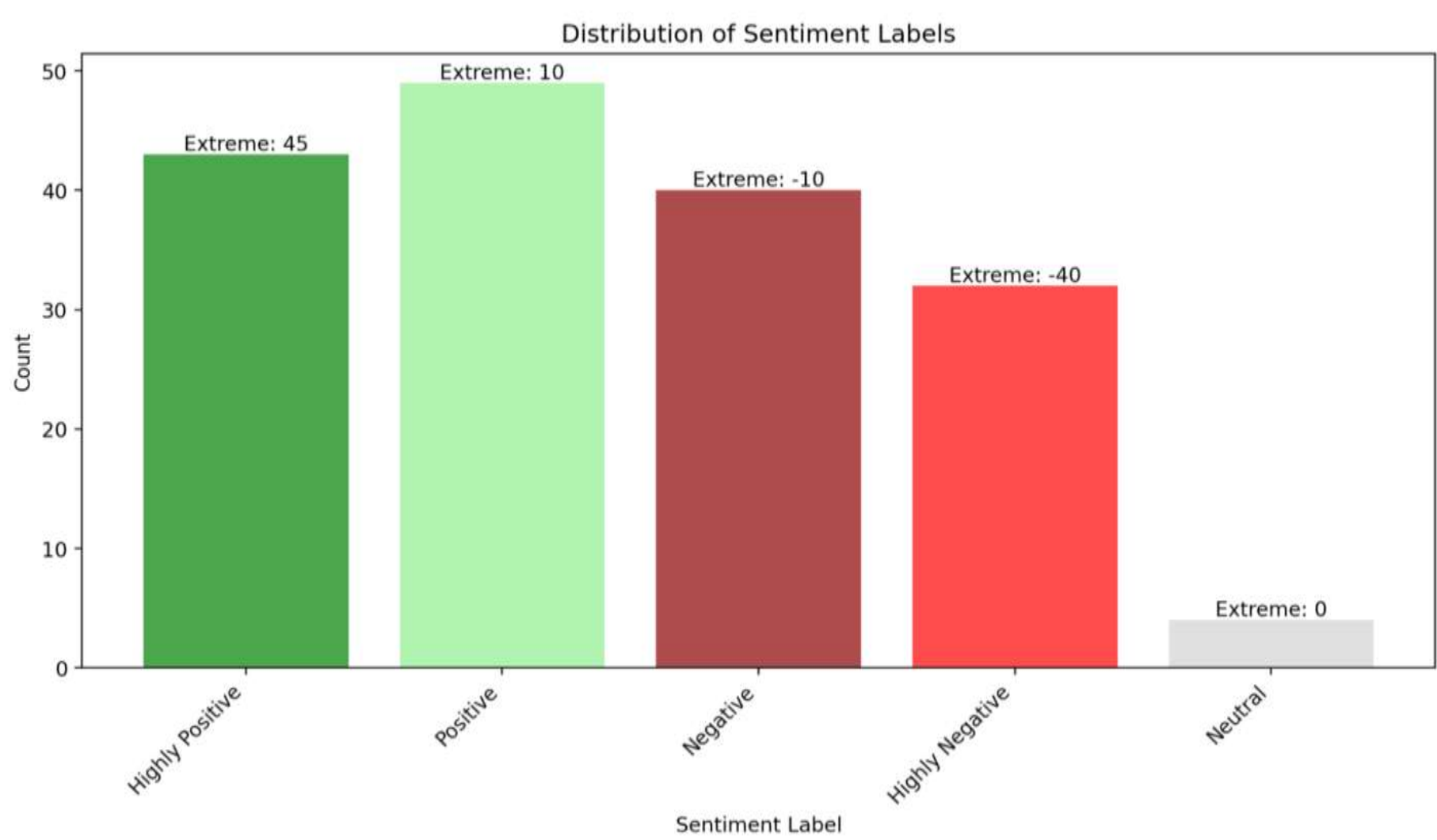


Fig 4: Distribution of Sentiment Labels with Extreme Scores

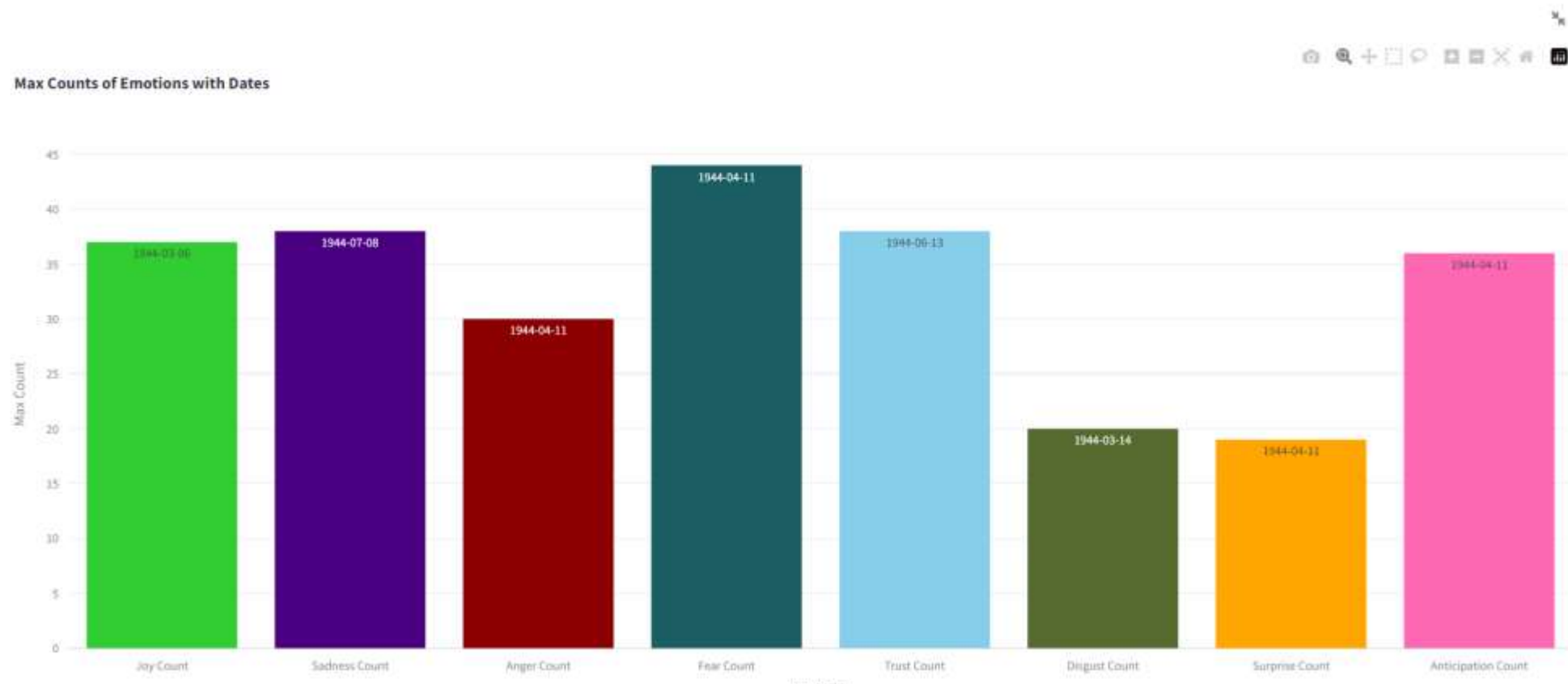


Fig 5: Highest Count of Each Emotion with Date

## Conclusion

Examining Anne Frank's diary sentimentally brings out the way that a young girl was emotionally going through difficult times in history. By means of close reading and visualizing, we find her emotion changes from complete depression to little episodes of optimism. The study adds depth to our knowledge of Anne's survival and resilience and provides insights into the human experience amid genocide.

## QR Code for Streamlit app & References

