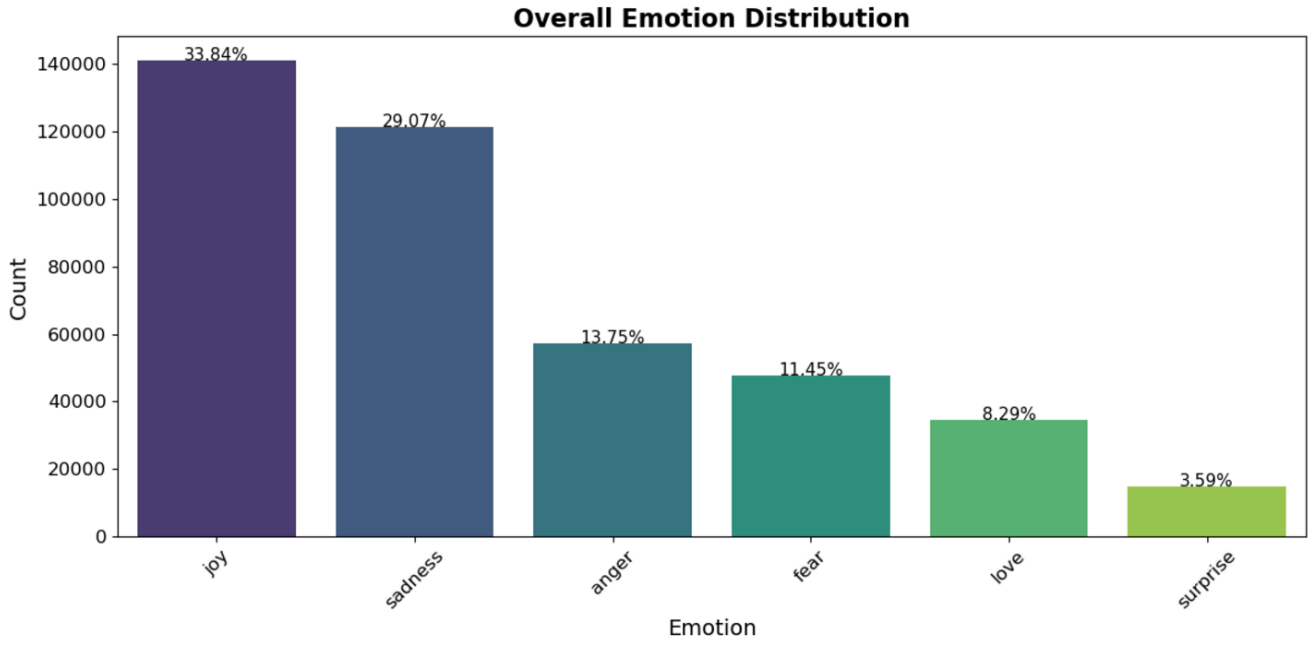
Text Data Visualizations

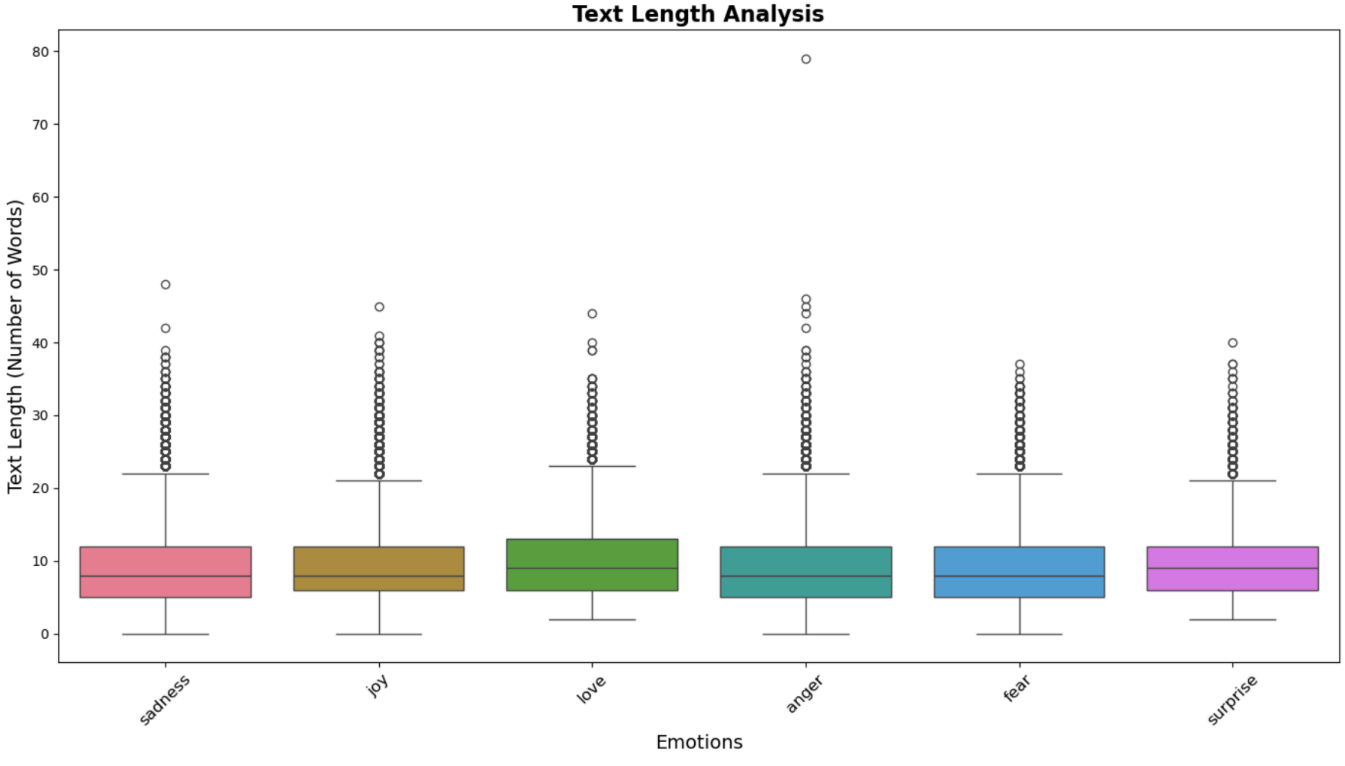
Distribution of Emotions :



Insights:

This graph actually shows how the emotions are distributed across the text dataset, overall I can see joy seems to be the most dominating emotion (33.84%) out of all the other emotions and Sadness (29.07%) is the next most dominant one there seems to be a skew towards these two emotions, people are expressing these emotions extremely in their text. On the whole, the negative emotions (sadness, anger, and fear) together are making up about 54% of the dataset, which I guess is quite revealing about how people express themselves online. Whereas, the other emotions like surprise and love are the least represented emotions overall they are making up only about 12% combined together.

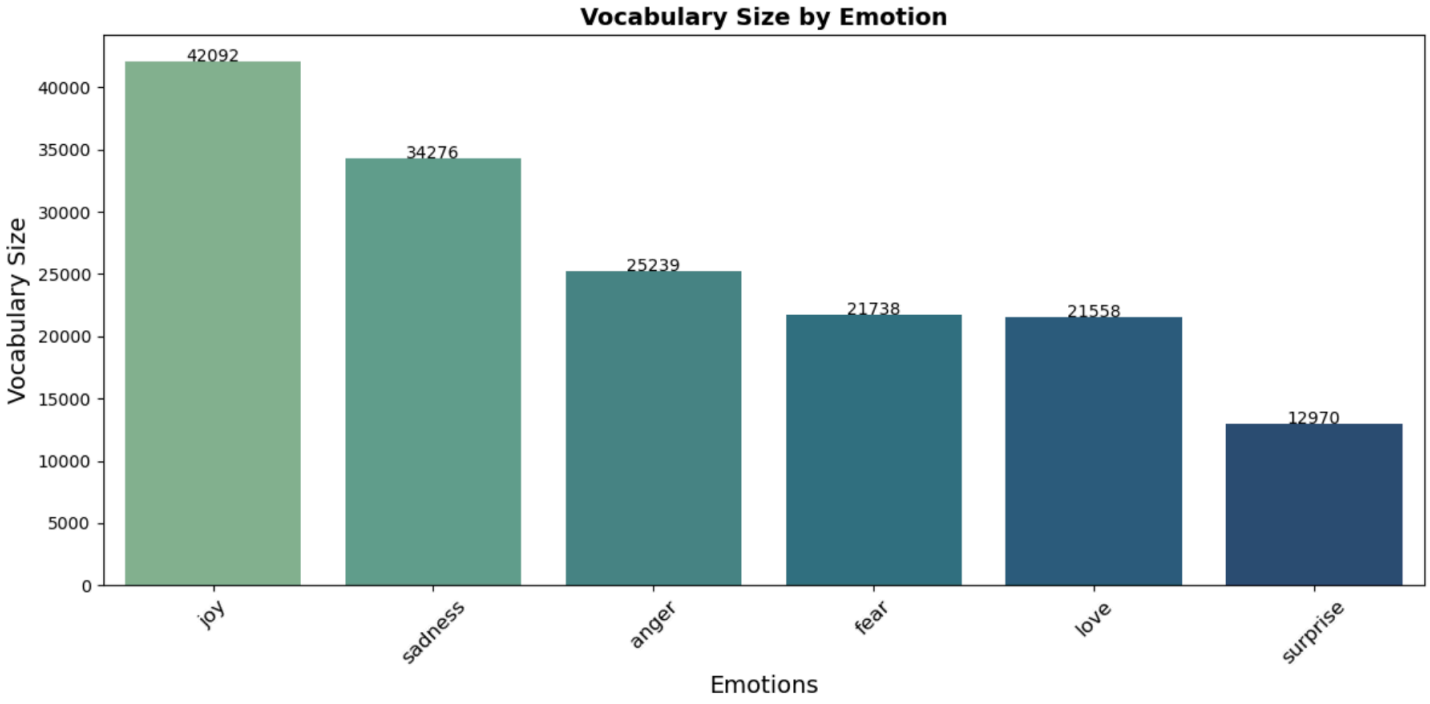
Text Length Distribution:



Insights :

This box plot basically shows the distribution of the length of the emotions expressed through texts, most of the emotional texts seems to be relatively concise, they are ranging between 5-15 words only. Looking at the medians I can see that people are expressing their emotions quite similarly in terms of length, regardless of whether they're happy, sad, or angry. I can also see there are some outliers present people are occasionally writing longer emotional texts and these seems to appear across all the emotions, showing that only when people feel strongly about something they might write more extensively, regardless of the type of emotion they're expressing.

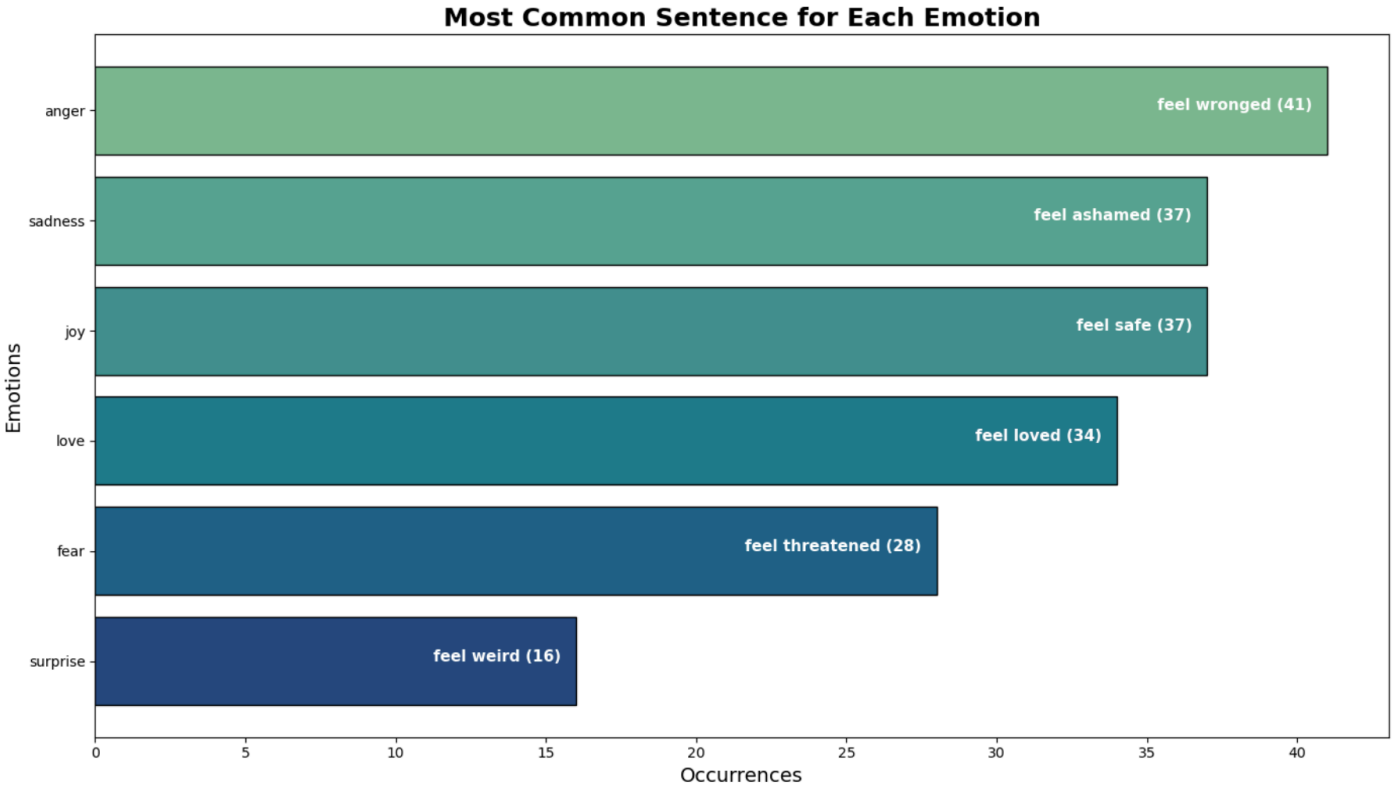
Distribution of Vocabulary Size Across Emotions:



Insights:

This bar chart basically displays the distribution of the number of vocabulary used for various emotions in the dataset. Overall, the vocabulary size gradually decreases from positive to negative emotions. From what I can see the emotion joy has the highest vocabulary count (42,092 unique words), followed by sadness with 34,276 words. The other emotions like anger (25,239), fear (21,738), and love (21,558) show somewhat similar vocabulary sizes, suggesting people use comparable ranges of words to express these feelings. On the other hand, surprise has the smallest vocabulary at 12,970 words, indicating people use more limited or specific phrase when expressing surprise. This pattern is showing that the positive emotions, particularly joy, tend to elicit more diverse vocabulary usage in textual expressions.

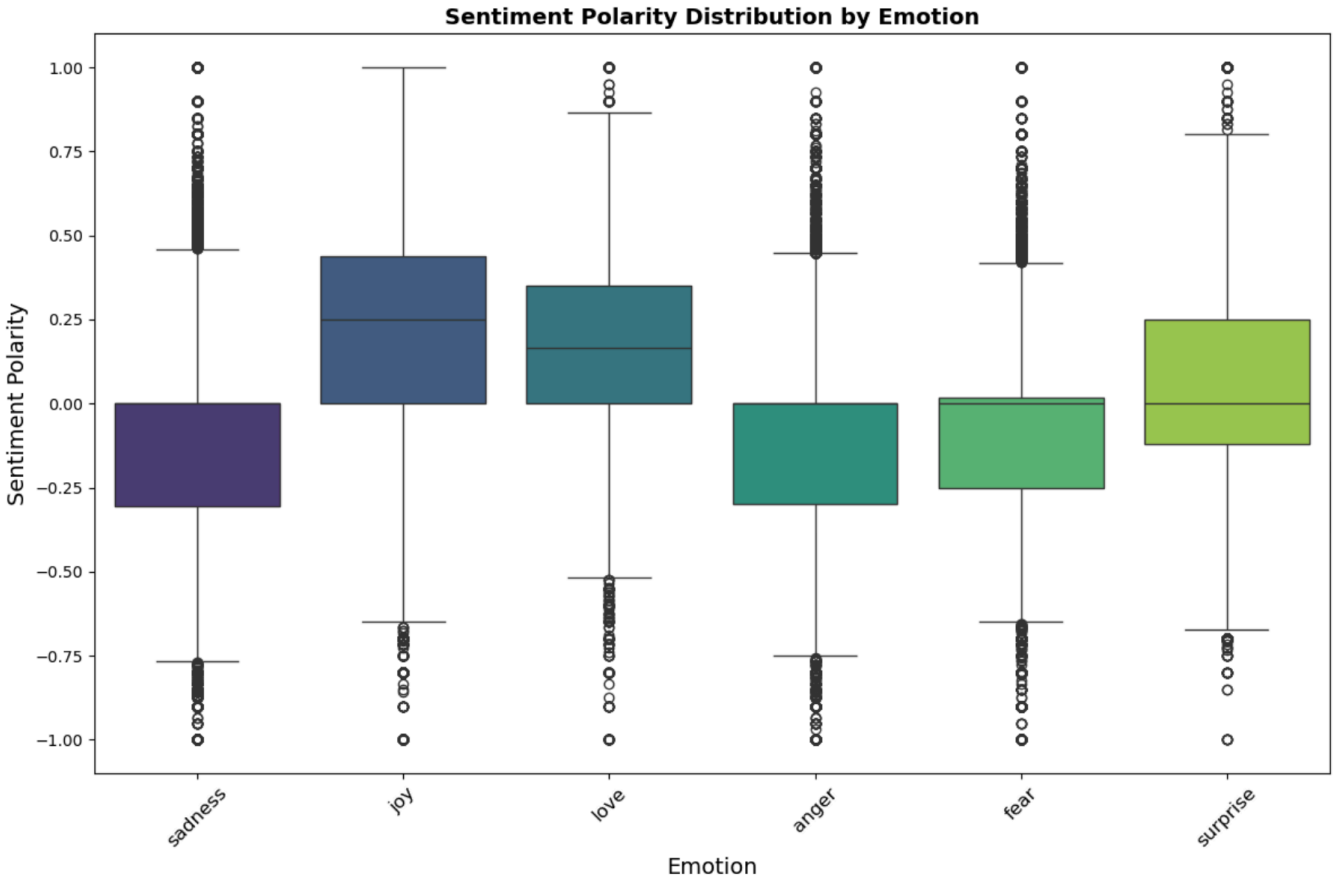
Most Frequently used complete Expressions by Emotion:



Insights :

This bar chart actually displays the most common expressions used for each emotion in our text dataset, and its quite interesting to see how people are articulating their different feelings, I can see the phrase "feel wronged" is used more commonly when people are angry people have connected anger with injustice and this phrase is the most dominating out of all the phrases expressed by people in their for different emotions. Whereas, when looking at sadness and joy these expressions come next people have expressed them pretty obviously the most common phrase used by people in the text when they were sad seems to be "feel ashamed" and on the other hand for joy its "feel safe". While for the rest of the emotions like love, fear and surprise people have used the direct phrases like feel loved, feel threatened, feel weird respectively more common.

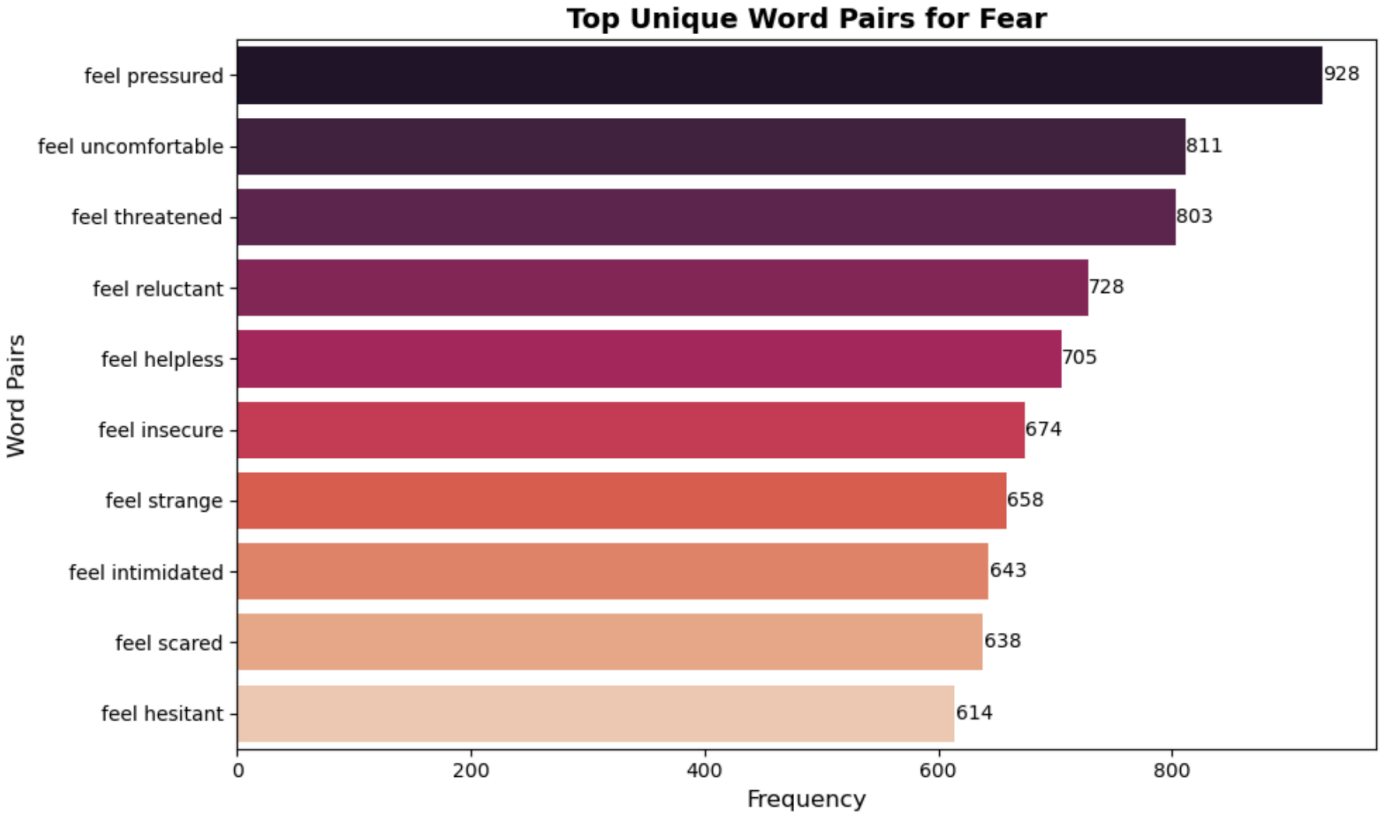
Sentiment Polarity Distribution Across Different Emotions :

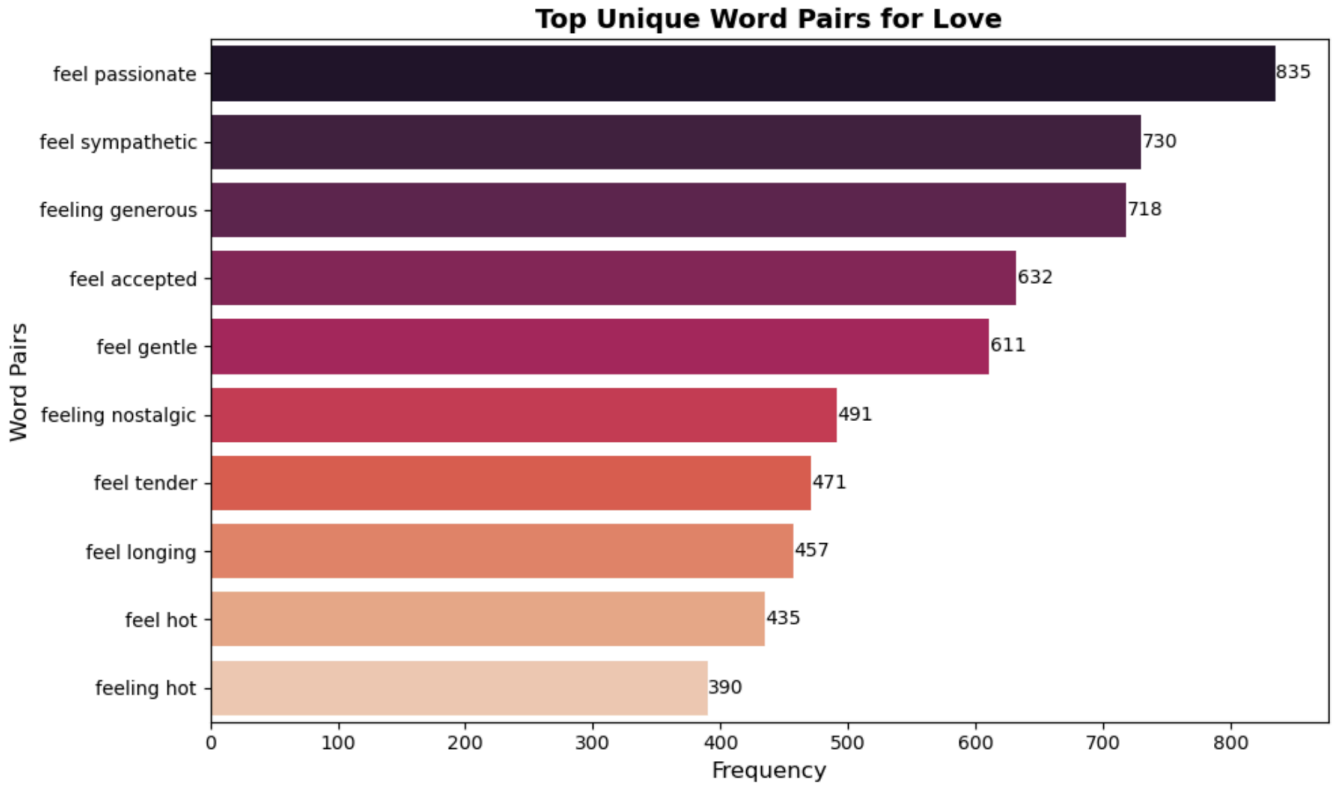


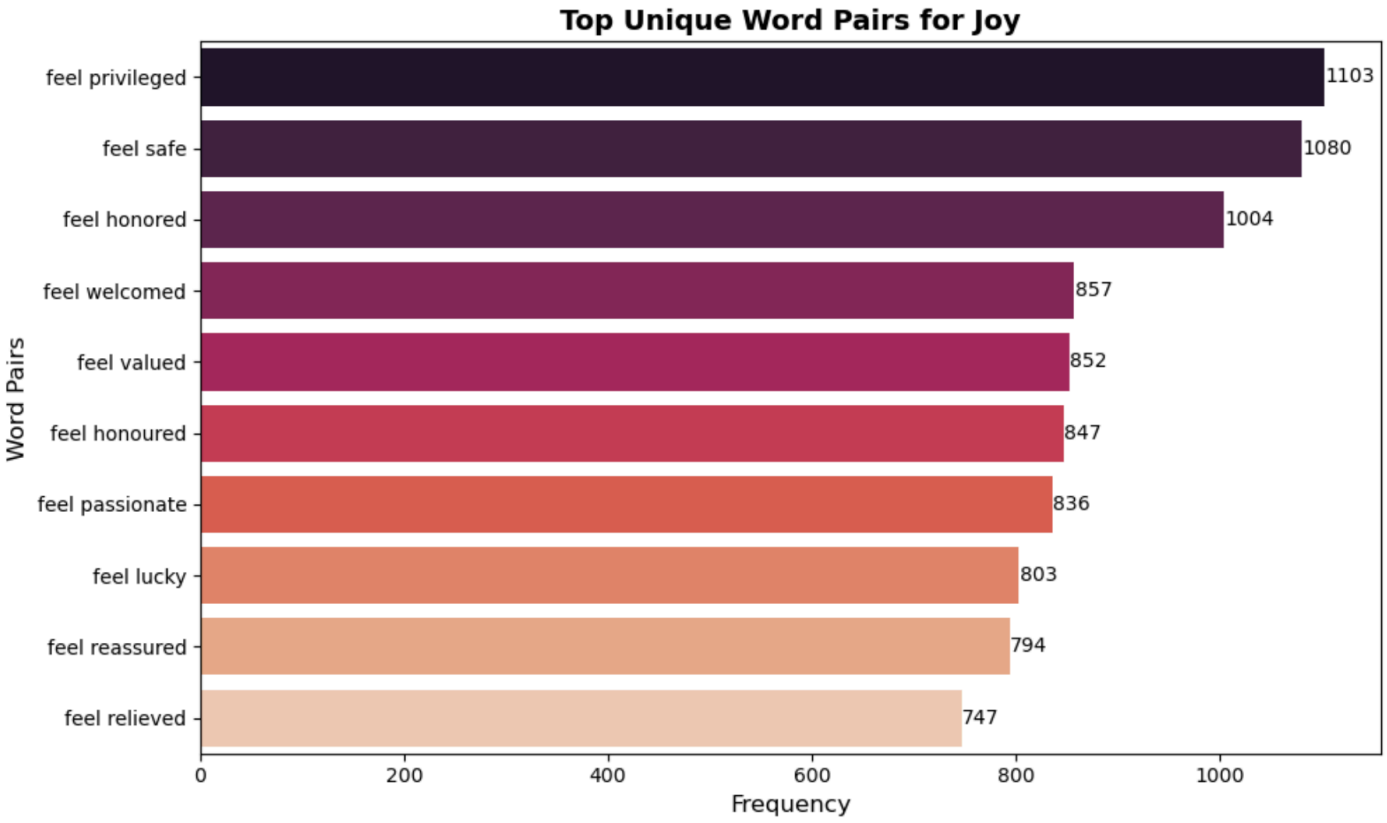
Insights :

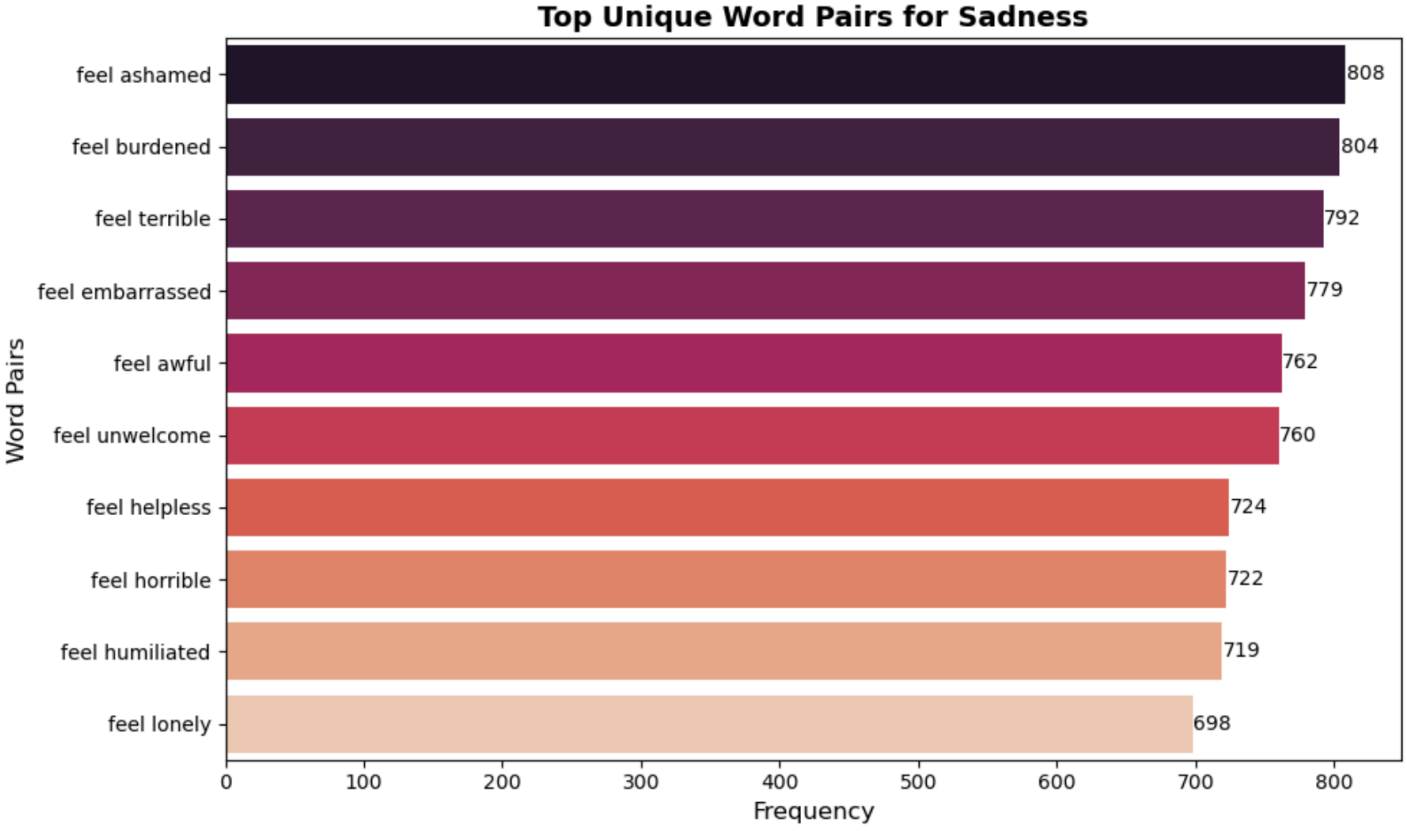
This box plot shows how sentiment polarity is distributed across various emotions, I can see that texts that express joy and love are having more positive sentiment scores there median's around 0.25 to 0.45, aligning with their positive emotional nature. While, sadness shows the most negative sentiment distribution its median is around -0.25, similarly anger and fear are also showing negative patterns but with varying distributions. Surprise on the other hand is showing a near-neutral median its indicating that it can be expressed with both positive and negative sentiments depending on context. And, pretty much all the emotions are displaying outliers at both the extremes +1.0 and -1.0. From the substantial overlap in the distributions I can see that while sentiment analysis helps to identify emotional content, it might alone not be sufficient to accurately distinguish between the specific emotions.

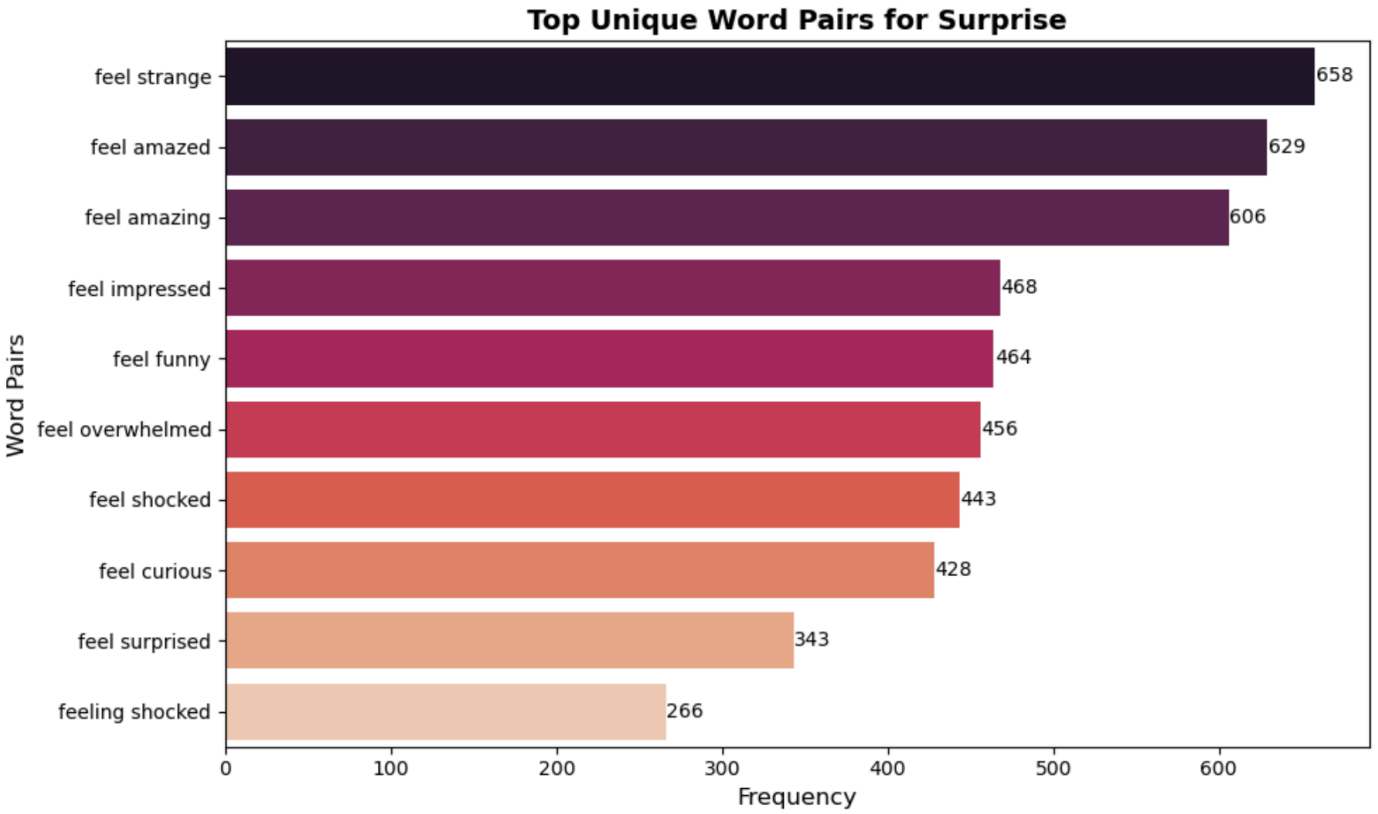
Top 10 Unique Word Pairs for Each Emotion:

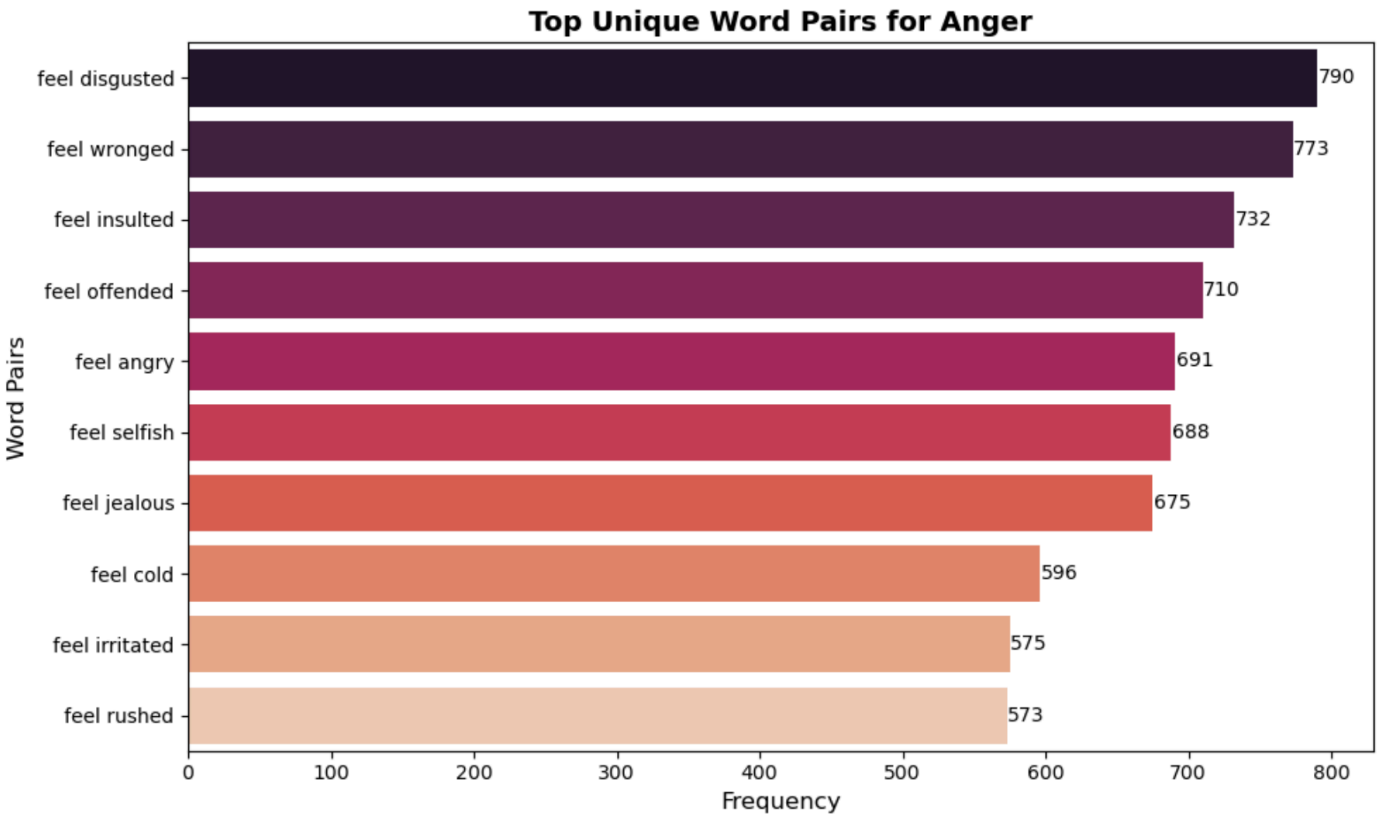





Top Unique Word Pairs for Joy
feel privileged
feel safe
feel honored
feel welcomed
feel valued
feel honoured
feel passionate
feel lucky
feel reassured
feel relieved
1103
1080
1004
857
852
847
836
803
794
747
Word Pairs
0
200
400
600
800
1000
Frequency


Top Unique Word Pairs for Sadness
feel ashamed
feel burdened
feel terrible
feel embarrassed
feel awful
feel unwelcome
feel helpless
feel horrible
feel humiliated
feel lonely
808
804
792
779
762
760
724
722
719
698
Word Pairs
0
100
200
300
400
500
600
700
800
Frequency

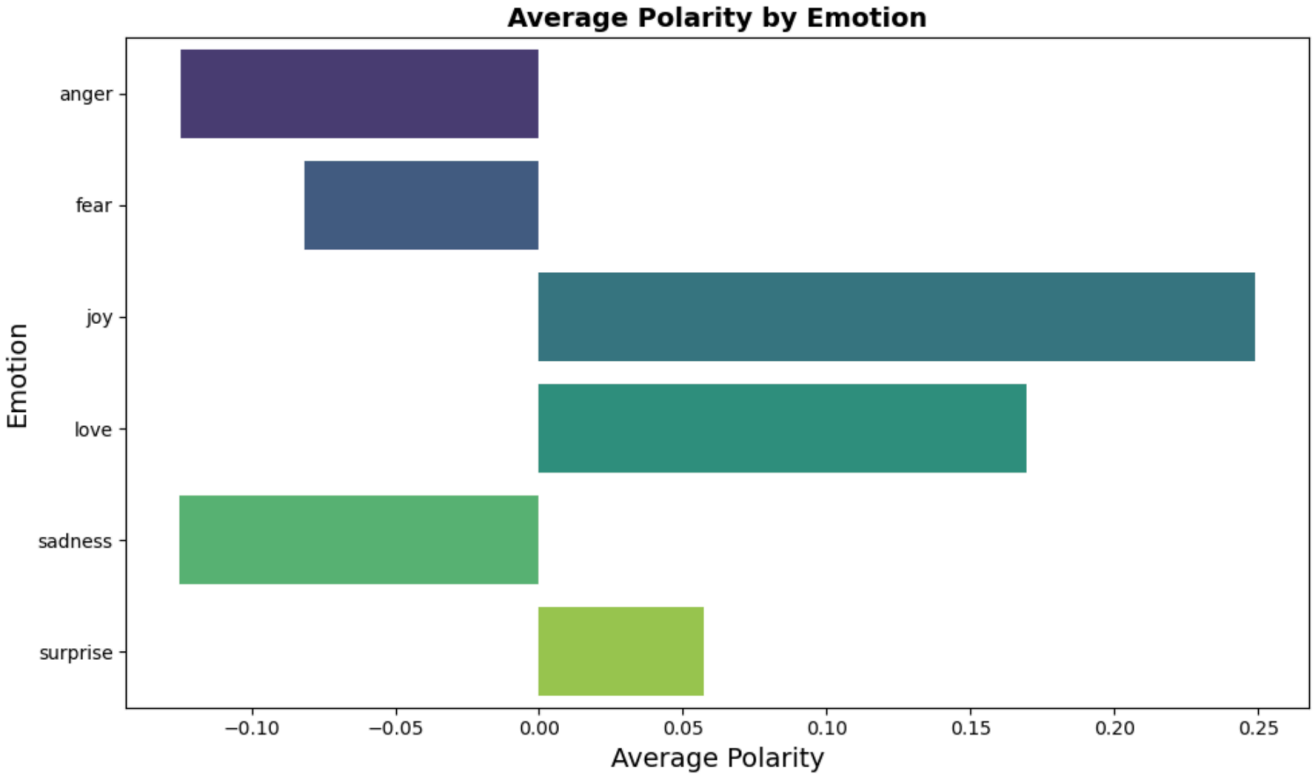

Top Unique Word Pairs for Surprise
feel strange
feel amazed
feel amazing
feel impressed
feel funny
feel overwhelmed
feel shocked
feel curious
feel surprised
feeling shocked
658
629
606
468
464
456
443
428
343
266
Word Pairs
0
100
200
300
400
500
600
Frequency


Top Unique Word Pairs for Anger
feel disgusted
feel wronged
feel insulted
feel offended
feel angry
feel selfish
feel jealous
feel cold
feel irritated
feel rushed
790
773
732
710
691
688
675
596
575
573
Word Pairs
0
100
200
300
400
500
600
700
800
Frequency

Insights :

These bar plots show the unique word pairs used most commonly to express different emotions in our text dataset, giving insights into how people verbalize their feelings. For joy, expressions like "feel privileged" (1103) and "feel safe" (1080) dominate showing positive states of well-being and security, while sadness I can see is characterized by self-reflective negative expressions like "feel ashamed" (808) and "feel burdened" (804). And, Fear expressions center mostly around vulnerability, with "feel pressured" (928) and "feel uncomfortable" (811) being the most prevalent, and love shows a range of empathetic and passionate expressions, like "feel passionate" (835) and "feel sympathetic" (730). On the other hand, I noticed that anger text expressions are predominantly reflecting perceived injustice and offense by phrases like "feel disgusted" (790) and "feel wronged" (773), while surprise, though having lower frequency counts overall, is actually showing diverse reactions ranging from "feel strange" (658) to "feel amazed" (629). This distribution of word pairs not only reveals the most common ways people express each emotion but also provides valuable insights into the linguistic patterns and emotional vocabulary used.

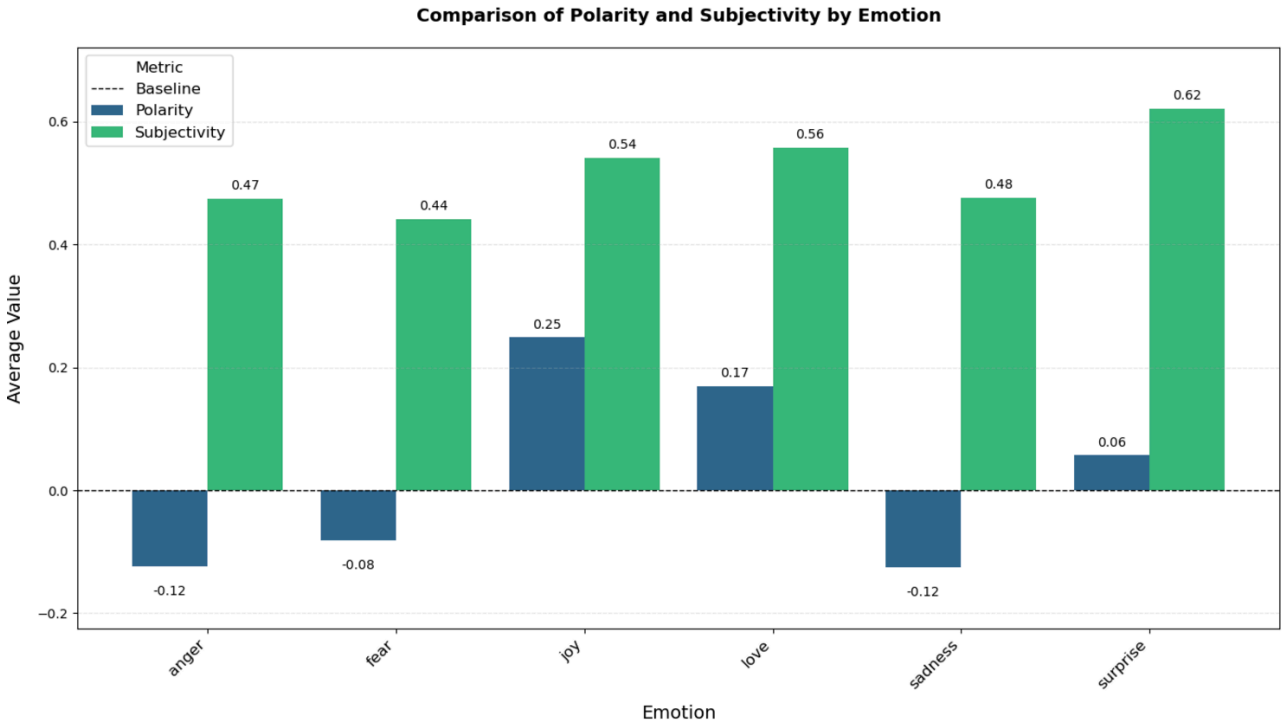
Emotion-Wise Average Sentiment Polarity Analysis :



Insights :

This plot actually shows the relationship between different emotions and their average polarity scores, measured on a scale from approximately -0.10 to 0.25. The length and direction of each bar show the polarity strength and whether it's positive or negative. On the positive side, joy and love seems be having the strongest positive associations, while anger and fear are showing notably negative polarities. This plot also shows how some emotions consistently align with positive sentiments (joy, love), while others like anger, fear, sadness toward negative sentiments, whereas surprise I can see is maintaining a relatively neutral position with a slight positive lean.

Comparing Polarity and Subjectivity Across Emotions :



Insights :

This plot shows the relationship between polarity and subjectivity across the different emotions. The polarity scores actually clearly distinguish between the positive and negative emotions, joy (0.25) and love (0.17) are showing positive values, while anger, sadness (-0.12) each, and fear (-0.08) have negative values. But, interestingly, I can see subjectivity maintains consistently high levels across all emotions ranging from 0.44 to 0.62, and surprise is having the highest subjectivity (0.62), showing that emotional expressions are subjective regardless of its positive or negative nature. The negative emotions on the other hand are showing a contrast, combining low polarity with high subjectivity scores, showing that negative feelings can be strongly subjective despite their negative orientation. Whereas for surprise there's slightly positive polarity (0.06) with the highest subjectivity score (0.62), it's both the most subjectively expressed emotion and slightly positive in nature.