Emotions in Music: A Comparative Sentiment Analysis of Popular Songs Across Countries and Markets

MATTEO DEL PRATO ALESSANDRO AMANDONICO
matteodp@kth.se | aleama @kth.se

January 17, 2025

Abstract

The study compares the sentiments of top-charting songs across different markets and nations, exploring the emotional aspects of popular music. It investigates how the sentiment expressed in song lyrics varies depending on the country and historical context, providing insights into the factors contributing to success in specific markets. The emotional content of song lyrics has been quantified using two different sentiment analysis tools such as VADER and TextBlob to provide a more robust representation and analysis of sentiment values in songs. Results were then compared with happiness metrics from the World Happiness Report to find patterns and correlations. The findings reveal distinct sentiment patterns: songs from European nations display higher instances of negative sentiment, while those from Latin America and North America lean towards positivity or neutrality. However, the correlation between national happiness scores and song sentiment is inconsistent, highlighting the limitations of current sentiment analysis tools in capturing cultural and emotional nuances.

Contents

List of Acronyms and Abbreviations

AR: Argentina
BR: Brazil
IT: Italy
ME: Mexico
SW: Sweden
US: United States of America
VADER: Valence Aware Dictionary and sEntiment Reasoner
NLTK: Natural Language Toolkit
GDP: Gross Domestic Product
API: Application Programming Interface
NLP: Application Programming Interface

1 Introduction

This research investigates geographical trends in music preferences through sentiment analysis of song lyrics, focusing on understanding how their emotional tone reflects cultural, social, and economic factors. By examining sentiment patterns in top-charting songs from various countries, the study aims to uncover how external influences - such as political unrest or economic stability - shape the emotional content of popular music. To do so, we are employing sentiment analysis: a natural language processing (NLP) technique traditionally used in domains like social media and customer feedback analysis. In this way, we can offer insights into global musical preferences but also evaluate the adaptability of machine learning techniques to less structured and more artistic forms of text, potentially expanding the scope of sentiment analysis in creative industries.

1.1 Previous Work

This section reviews the existing literature, exploring previous work on sentiment analysis of song lyrics, the role of music in society, and the challenges of applying NLP to creative texts.

1.1.1 Sentiment Analysis in Music Lyrics

The application of sentiment analysis to lyrics has gained interest in recent years: Makris [1] demonstrated how sentiment analysis could be applied to music lyrics by analyzing the works of a Greek rock band. His project focused on using sentiment analysis tools like VADER and TextBlob to assess the emotional tone of their lyrics. Venditama [2] expanded on the application of sentiment analysis by scraping a large dataset of Indonesian song lyrics, findings highlighted recurrent themes of love, relationships, and betrayal, with melancholic tones. Venditama's work also underscored the importance of language-specific preprocessing, such as cleaning stopwords and employing tools like Sastrawi, to improve the accuracy of sentiment analysis in diverse languages.

1.1.2 Music as a Reflection for Society

Beyond individual emotional expression, music is often seen as a reflection of society's collective mood. In a recent study, Edwards [3], argued that music, particularly pop music, captures the emotional atmosphere of a given period due to war, economic hardship, or social progress. Similarly, Yalcinkaya [4] conducted a study that analysed the sentiment of song lyrics in relation to sociopolitical factors, finding that periods of social unrest were associated with more aggressive lyrics. Finally, Venditama's analysis [2] of Indonesian song lyrics further supported this idea by identifying themes of love, betrayal, and sadness. His study highlighted how personal emotional experiences in music often parallel broader societal issues. These findings underscore the complex interplay between music and societal emotions, yet they leave unanswered questions about how these dynamics vary across cultures and geographic contexts. For example, while regional studies exist, comprehensive cross-cultural analyses of sentiment in popular music remain unexplored.

1.1.3 Music and Happiness

The relationship between music and societal happiness has also been explored throughout the years: Benetos [5] explored how national happiness in the UK could be inferred from musical sentiment, suggesting that countries with a higher proportion of positive lyrics in popular music might experience higher overall happiness. Li [6] similarly examined how the emotional tone of music contributes to individual and societal well-being, arguing that positive songs promote happiness and emotional well-being.

1.1.4 The Challenges of Applying NLP to Music Lyrics

Antonio et al. and Jose et al [7] discussed the limitations of traditional NLP techniques when applied to lyrics, emphasising that music often deviates from conventional grammatical structures. Metaphors, slang, cultural references, and poetic expressions in lyrics complicate the application of sentiment analysis tools.

1.1.5 Contribution

While existing literature establishes music as a societal mirror and highlights the potential of sentiment analysis tools, there is a lack of comprehensive, cross-cultural studies integrating these perspectives. This research aims to bridge this gap by analyzing how lyrical sentiments correlate with national happiness across diverse countries and cultural contexts.

1.2 Research questions

The research aims to address the following key questions:

1. Are there emotional tone differences in top-charting songs between countries?
2. How do lyrical sentiment trends correlate with happiness factors between countries and regions?

1.3 Hypotheses

The study will propose three hypotheses:

1. Lyrics of popular songs from politically or economically troubled nations will likely show more negative sentiments than those from stable countries.
2. Mainstream national songs will reflect the population's sentiments during the specific period.
3. Songs that reflect strong emotion of public sentiment are more likely to achieve popularity than neutral ones.

2 Method

2.1 Overview

In this section, we describe the methodology used to investigate the emotional value of song lyrics across countries and regions. In particular, our study combined: data collection, the sentiment analysis itself, and a comparative analysis to uncover trends and differences. The countries we took into consideration for this study were chosen to represent three different macro-regions: North America, South America and Europe. Each region has two representative states, each of them listed as follows.

Region	Countries (Abbreviations)
North America	USA (US) Mexico (ME)
South America	Argentina (AR) Brazil (BR)
Europe	Sweden (SW) Italy (IT)

Table 1: Regions and their representative countries

The reason why those countries and regions have been chosen depends solely on difficulties regarding languages and datasets. For example, the unavailability of some Asian and African countries' top chart playlists in the Spotify catalogue [8] prevented us from analysing that data. Furthermore, during the process of sentiment analysis, we needed to download datasets of stop words for each different language too, but the NLTK (Natural Language Toolkit) we used was missing many languages [9]. Figure 4 in Appendix A shows the list of the available languages we could use. After making these considerations we decided to focus on a few languages (Italian, Swedish, English, Spanish and Portuguese) and we excluded from our study some regions due to lack of data (Asia, Africa, Oceania). We began by gathering song data from Spotify's country top charts and retrieving each song's corresponding lyrics from Genius and Musixmatch databases. We then used two different natural language processing tools, VADER and TextBlob, to quantify sentiment and analyse both positive and negative tones in textual data. Finally, we compared the results of sentiment analysis and looked for patterns across different cultural and geographic contexts, comparing them to the general happiness values of each country taken from the World Happiness Report database. The following subsections give more details on each stage of the methodology, highlighting our rationale and adjustments to ensure the findings' robustness and relevance.

2.2 Data Collection

Data for this study was collected using Spotify's API to identify the 50 top-charting songs in the selected countries, and save relevant details of each song such as title, author and album name in an Excel sheet for later use. The reasons behind choosing Spotify were its comprehensive catalog of music and its accessible API which allowed for efficient extraction of top-charting songs by country. After downloading country-specific chart's songs metadata, we extracted song lyrics for each song using easily accessible databases like Genius and Musixmatch; title and author were used as queries to find and download the corresponding lyrics. A Python script was utilised, automating the collection and storage of song metadata and lyrics into Excel sheets for further analysis. Tracks were then filtered to exclude instrumental pieces or incomplete metadata entries. Emoji and special symbols removal and normalisation were applied to ensure data uniformity and safe search of the lyrics.

2.3 Sentiment Analysis

After collecting all the data on the songs, preprocessed lyrics were analysed using the VADER and TextBlob sentiment analysis tools. Firstly, the text was broken into smaller units, such as words or phrases, which can then be individually analyzed, then it was lemmatized, so words were reduced to their base or dictionary form (lemma) to ensure consistency, and finally text was cleaned of stop words (language-specific lists extended to include common lyrical filler words). Then, sentiment scores were generated (see Appendix B for detailed structure of scores) and an overall value of sentiment (positive or negative) was assigned to each song.

Open-source tools like VADER and TextBlob were selected for their ease of integration with the Python script, for their established reliability in sentiment analysis tasks and to minimize resource-intensive processes. Two different sentiment analysis tools were utilised to produce a more thorough representation and analysis of sentiment values for songs and to have more robustness for our findings. If on one hand VADER excels at analyzing short, informal texts like tweets or reviews by leveraging a lexicon of words with associated sentiment intensities, making it great for quick, context-sensitive sentiment scoring. TextBlob, on the other hand, provides a more general-purpose approach, combining rule-based sentiment analysis with natural language processing tools, but it can struggle with nuanced or slang-heavy language.

2.4 Comparative Analysis

Finally, we integrated lyrical sentiment data with external happiness metrics to analyse and compare sentiment trends across countries. The sentiment scores for the songs were aggregated by country to determine the average sentiment values, given by positive, negative, and neutral scores. These values were then compared with happiness rankings obtained from the World Happiness Report [10]. The happiness rankings take into consideration various aspects of a country's wellness, to see all the parameters taken into consideration see Appendix C. Data visualisation techniques, such as box plots and scatter plots, were then employed to illustrate the relationship between sentiment scores and happiness indices. Correlation analysis further explored these relationships, identifying patterns and trends that might suggest connections between a nation's happiness and the emotional tone of its popular music. Finally, clustering analysis was utilised to highlight and discover patterns between continents and countries values of emotional tone.

3 Results and Analysis

3.1 VADER: Analysis of Results

VADER sentiment analysis (see Figure 6 in Appendix D) shows that positive sentiment is predominant, with ME and AR scoring highest. While European countries exhibit more negative sentiment, Southern American countries are the only ones to report neutral sentiment. As shown in Table 2 in Appendix D, sentiment varies by region: Southern American countries and ME have a positive interquartile range, while the US and Europe show mixed sentiments (see Figure 7 in Appendix D). Notably, average compound scores for the US and BR are similar, whereas IT has a significantly lower score (see Figure 8 in Appendix D). Country clusters based on mean compound scores (see Figure 9 in Appendix D) indicate that AR and ME have the highest positive sentiment. BR and the US show mixed, positively leaning sentiments, while IT and SW display slightly positive scores.

3.2 TextBlob: Analysis of Results

According to TextBlob's sentiment analysis, neutral sentiment is predominant in most countries, with the US and SW showing prevailing positive sentiment and only IT exhibiting negative sentiment (see Figure 10). The mean and median Polarity scores are close to 0 across countries (see Table 3 in Appendix E), reflecting this neutral trend, despite the presence of both positive and negative sentiments. As shown in Figure 11 in Appendix E, AR, BR, and ME display consistent neutral sentiment with narrow interquartile ranges and evenly distributed outliers, while IT, SW, and the US have wider ranges but fewer outliers. Interestingly, SW has a high concentration of negative sentiment but a relatively high mean polarity score (see Figure 13). The US stands out for its high level of personal opinion, while European nations tend to contain more emotional content compared to AR,BR and ME, although scoring below a 0.5 threshold (see Table 4 in Appendix E). The distribution of Subjectivity scores (see Figure 12 in Appendix E) shows that IT has the widest range, indicating that emotional tones can be both positive and negative. When comparing clusters of countries based on mean Polarity scores (see Figure 14 in Appendix E) and those based on mean Subjectivity scores (see Figure 15 in Appendix E), results suggest that countries with higher subjectivity scores do not necessarily exhibit higher polarity scores. This suggests that higher subjective expression does not always correspond with strong sentiment polarity, indicating that subjective content can range from neutral to strongly opinionated sentiments.

3.3 Comparing Sentiment Scores

Figure 1 reveals how VADER frequently classifies more sentiments as strongly positive or negative, whereas TextBlob has a greater reliance on neutral classifications. VADER's effectiveness in capturing pronounced emotional expressions opposed to TextBlob's conservative attribution is evident when considering countries such as ME, AR and BR. Score distribution in Figure 2 shows how VADER is able to detect a broader range of sentiment values - including extreme positives and negatives - thus resulting in higher variability and wider interquartile ranges. In contrast, TextBlob generally produces more moderate and concentrated polarity scores, indicating a tendency for less extreme sentiment evaluations. When comparing median values, these show significant differences with higher values for VADER and lower values, mostly around 0, for TextBlob.

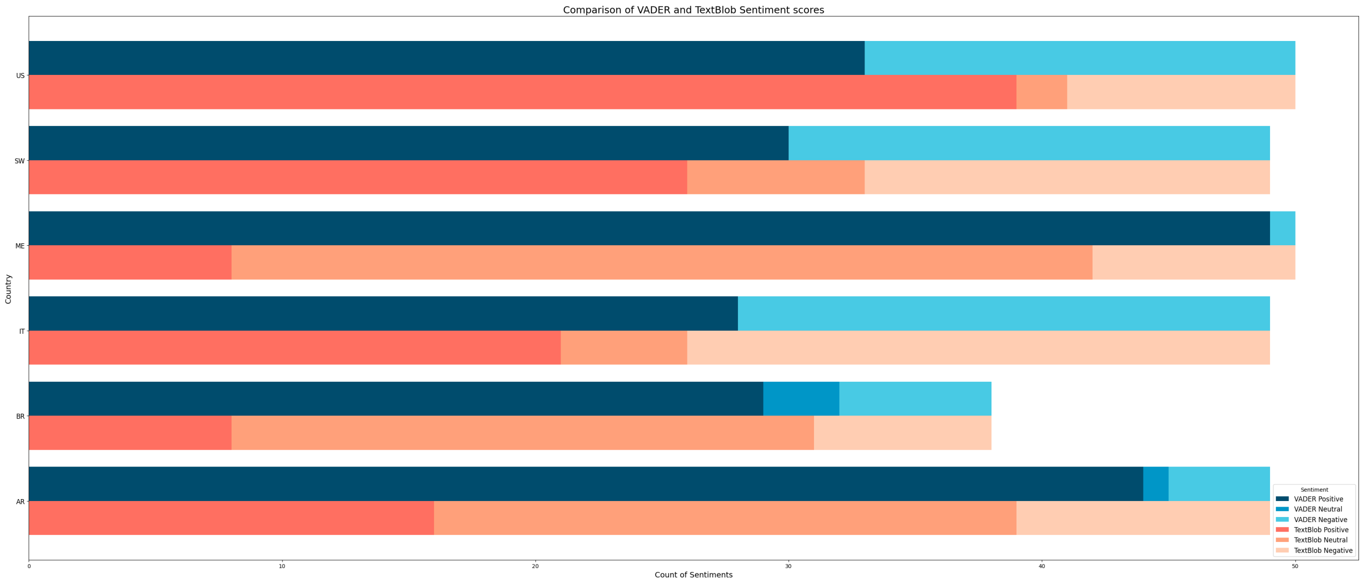

Figure 1: Sentiment score comparison per country across VADER and TextBlob

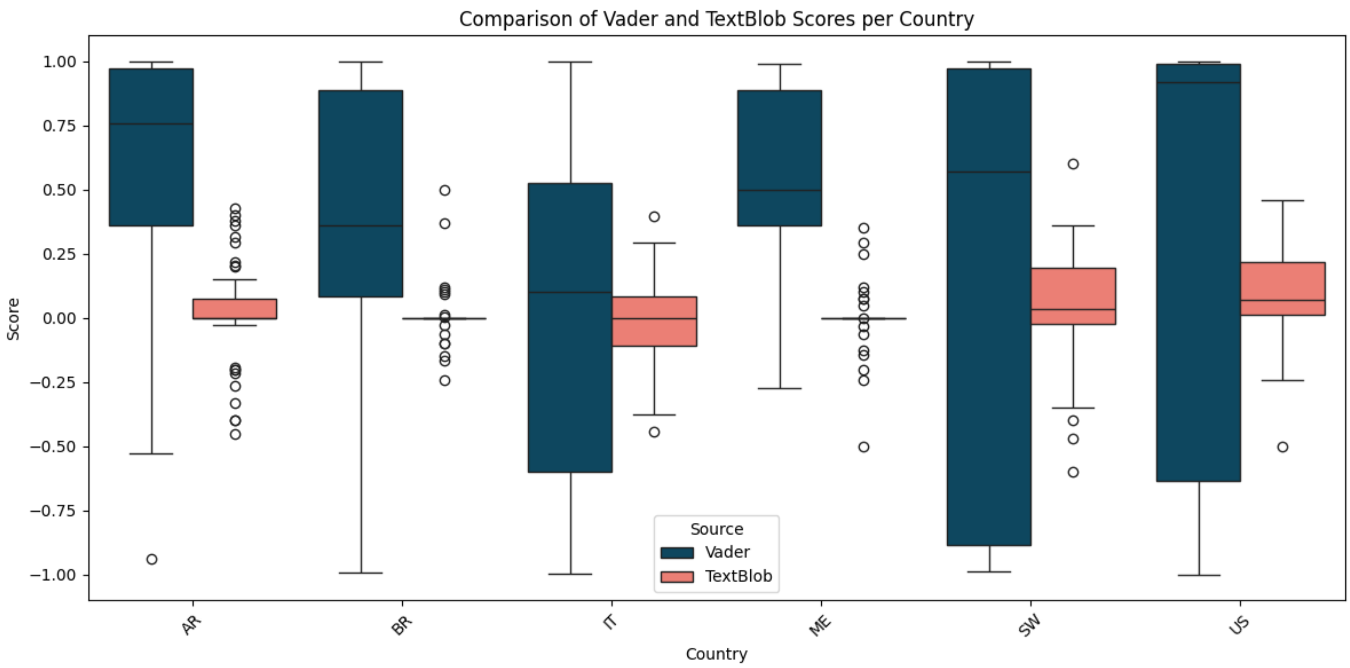

Figure 2: VADER's mean Polarity score and TextBlob's mean Subjectivity score distribution comparison per Country

3.4 Comparing Sentiment and Happiness Scores

According to happiness scores, SW ranks as the happiest country, followed by North American, IT, and Southern American countries (see Table 5 in Appendix F). VADER classification indicates a non-linear relationship with happiness rankings, resembling results only for IT (see Figure 16 in Appendix F), while TextBlob's mean polarity scores show little variation across happiness levels, suggesting a weaker alignment with observed happiness, especially in European countries and the US (see Figure 17 in Appendix F). Despite some alignment, significant discrepancies exist between the two classifications. Figure 5 shows that TextBlob's Polarity aligns more closely with perceived happiness than VADER's Compound score. TextBlob also displays higher positive correlations with happiness sub-indices, like *GDP per capita*, *Generosity*, and *Social support*, while VADER captures more negative sentiment dynamics, showing stronger negative correlations with factors like *Healthy life expectancy* and *Generosity*.

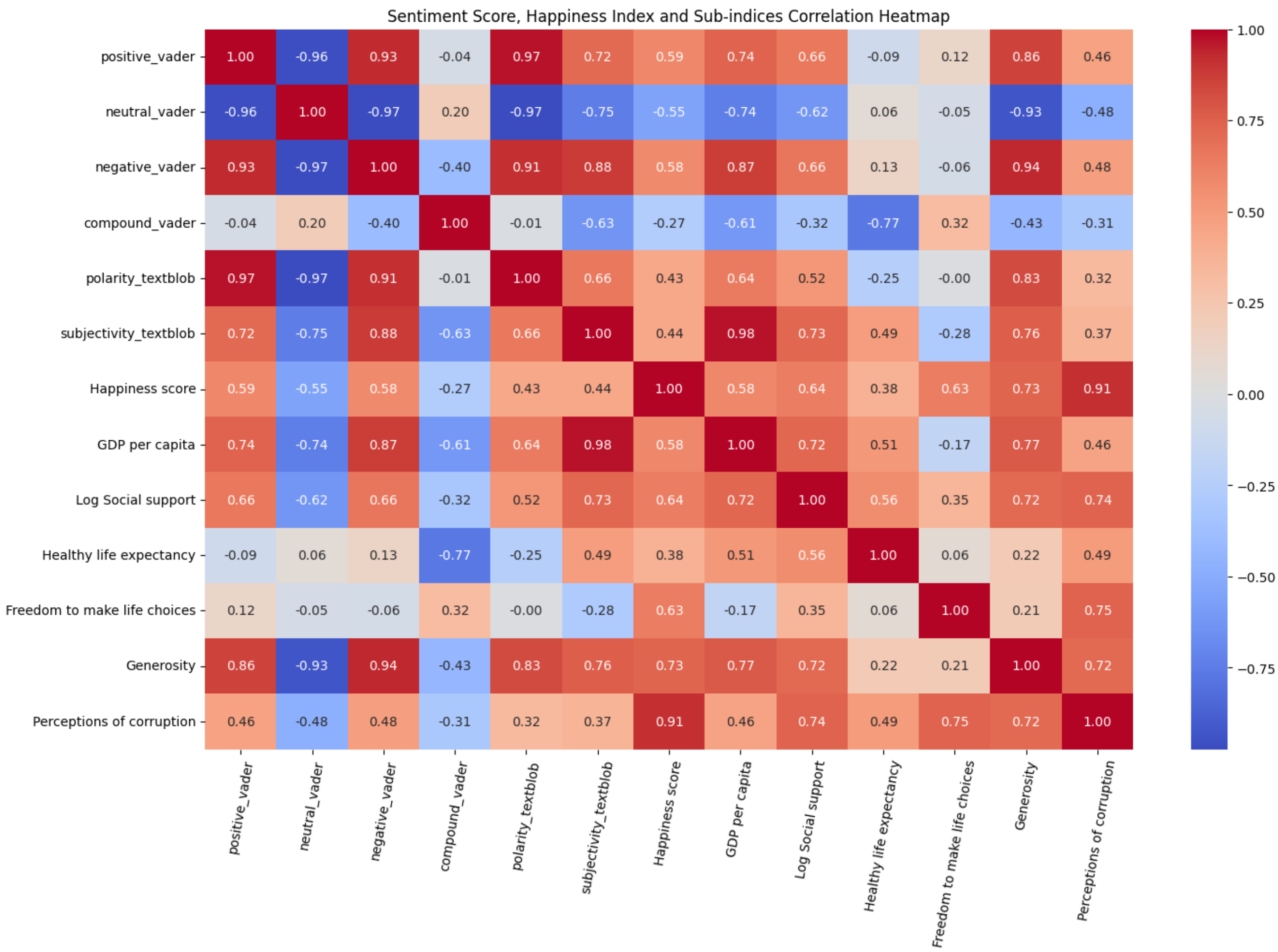

Figure 3: Correlation analysis of Sentiment metrics, Happiness index and sub-indices

4 Discussion

The study outputs address the two research questions:

1. Are there emotional tone differences in top-charting songs between countries?

   Emotional tone differences in top-charting songs do exist between countries, as shown in Figure 2. Positive sentiments dominate globally but vary in intensity and balance across regions, whereas negative sentiments are more prominent in European countries. This finds grounding in research:

cultural contexts significantly shape musical emotion, with tonal structures and lyrical expressions reflecting regional emotional frameworks [11] which, in this case, may correspond to continents or countries. These findings align with the notion of universal recognition of emotions in music, yet with culturally nuanced interpretations [12]. As shown by both VADER and TextBlob, the US and SW exhibit a general positive trend but with high variability of sentiment. AR and ME also display general positive sentiment but with less variability, suggesting a more uniform emotional tone that resonates with collectivist cultural traits often observed in Latin America [13]. In contrast, BR leans more towards neutrality of sentiment without peaks, which may reflect cultural preferences for balanced emotional expressions in creative outputs [14]. On the other hand, both tools identified IT as having the most negative sentiment, which aligns with cultural tendencies toward heightened emotional expressiveness often observed in Mediterranean contexts [15]. Cultural and contextual factors appear to influence these trends significantly, with sentiment tools differing in sensitivity and interpretation. The variability in emotional tone, as observed across countries, reinforces the complex interplay of individual and societal emotions in music, making it both a reflection and a driver of collective sentiment [11, 15].

2. How do lyrical sentiment trends correlate with happiness factors between countries and regions?

   The correlation between sentiment trends and factors that contribute to happiness varies depending on the tool used for the analysis. The results of VADER reveal that all countries have a predominant positive sentiment (see Table 6 in Appendix G), probably because the VADER lexicon detects slang and abbreviations more effectively, while TextBlob is better suited for formal language usage [16]. As shown in Figure 21 in Appendix G, while the aggregated compound score weakly correlates with the Happiness Score, individual metrics capture specific sentiment dimensions, reflecting its broader sentiment average. More neutral sentiment is associated with lower happiness values, while there is some overlap between negative sentiment and happiness levels, potentially capturing nuances in sentiment expression. A stronger - although still moderate - correlation is present in TextBlob. Countries with a predominant positive sentiment match those with higher Happiness scores (see Figure 19 in Appendix G), while in countries with a predominant neutral sentiment, a higher percentage of neutrality corresponds to a higher happiness score (see Table 6 in Appendix G). Moreover, both *Polarity* and *Subjectivity* metrics demonstrate a stronger positive correlation than VADER's compound score.

The results from the study address the three initial hypotheses:

1. The lyrics of popular songs from politically or economically troubled countries are likely to show more negative sentiments than those from stable countries

   Contrary to the hypothesis, emotional weight in lyrics appears context-dependent and not exclusively tied to economic or political stability. Economically (see Figure 21 in Appendix G) and politically (see Figure 26 in Appendix G) troubled nations exhibit lower negative sentiment counts compared to economically prosperous nations, as indicated by both VADER (see Figure 22 and Figure 27 in Appendix G) and TextBlob (see Figure 23 and Figure 28 in Appendix G), thus contradicting the initial hypothesis. The VADER analysis accounts for this trend by identifying a higher positive tone in underprivileged clusters (see Figure 24 and Figure 29 in Appendix G), suggesting a coping or cultural mechanism to address adversity through creative outputs [17]. Findings could also indicate different cultural approaches to processing trauma and hardship through art, rather than simply measuring emotional well-being. In contrast, TextBlob suggests that economically and politically troubled nations express less intense sentiments overall, resulting in lower polarity values (see Figure 25 and Figure 30 in Appendix G), yet leaning towards neutrality rather than negativity. This pattern might be attributed to differences in how cultures express emotions: what seems like neutrality in Western sentiment analysis tools could actually reflect unique, culturally specific ways of communicating feelings [18]. More prosperous countries, however, convey more emotional weight in the lyrics, both positively and negatively. This may be influenced by cultural norms that promote more direct emotional expression, or by market dynamics that favour content with higher emotional intensity in these areas [19].

2. Mainstream national songs will reflect the population's sentiments during the specific period

   While there is little alignment, the reflection is inconsistent across countries and tools (see Figure 32 in Appendix G), revealing complex cultural patterns in emotional expression through music. SW has a high happiness score but only shows moderate positivity in song sentiments using VADER and TextBlob, which might reflect Nordic cultural tendencies toward emotional restraint in artistic expression rather than actual happiness levels [20]. The US's score aligns well with positive song sentiments, possibly due to its individualistic culture that encourages direct emotional expression and a commercial music industry that often favors overtly emotional content [21]. Meanwhile, countries with lower happiness scores exhibit mixed trends, suggesting that cultural approaches to processing and expressing emotions through art vary significantly across societies [22]. These inconsistencies suggest that factors beyond happiness may influence song sentiments, complicating the direct relationship between the two. The variation could be attributed to diverse cultural traditions in artistic expression, differing societal norms regarding emotional display, and varying roles of music across cultures—whether as entertainment, cultural preservation, or social commentary [19]. Moreover, the correlation analysis may be affected by measurement biases, as sentiment analysis tools developed in Western contexts might not effectively capture the nuanced ways different cultures encode emotional content in their lyrics [23].

3. Songs that reflect strong emotion of public sentiment are more likely to achieve popularity than neutral ones

   While VADER analysis supports this finding, TextBlob challenges the universality of the hypothesis (see Table 6 in Appendix G), highlighting the complexity of measuring emotional expression across different cultural contexts [24]. VADER only detects AR and BR as countries showcasing neutral sentiment, and even then, it is at a low percentage. In contrast, TextBlob reveals that neutral sentiment is predominant in three out of the six countries analyzed. Neutral tones are predominantly detected by TextBlob's subjectivity score as well, with all nations expressing objective content except for the US (see Table 4 in Appendix E). The distinctions between tools may reveal not only methodological variations but also fundamental differences in the ways emotions are encoded within different languages and cultural traditions [25]. What appears as neutral in one cultural context might carry significant emotional weight in another, and the tools' variable ability to identify these nuances suggests the need for more culturally sensitive approaches to sentiment analysis [26]. Furthermore, the correlation between sentiment scores and cultural factors may be complicated by variables such as translation effects, market dynamics, and diverse traditions of artistic expression across societies [27].

4.1 Limitations

The study's scope is limited by the availability of data, restricting the analysis to three regions. Expanding the data collection methods or integrating multiple data sources could enhance the generalizability and depth of the findings. Additionally, while VADER and TextBlob provide valuable insights, their training on social media text introduces challenges in interpreting the nuanced and metaphor-rich language of song lyrics. Future enhancements to these models, including training on domain-specific datasets, could improve their reliability for analyzing creative texts. The study relies on visualization and summary statistics to present findings: while these methods offer an accessible overview, the lack of inferential statistical techniques, such as ANOVA or t-tests, limits the statistical generalizability of the results. Furthermore, aggregating data by country averages may obscure individual variations or outliers within datasets, highlighting the need for more granular analyses in future research.

4.2 Ethical and Sustainability Considerations

The study adhered to the API usage policies of Spotify [8], Musixmatch [28], and Genius [29], ensuring that all data retrieval processes were ethical and compliant. As no user-specific data was collected, risks to individual privacy were effectively minimized. To further uphold ethical standards, algorithmic bias was mitigated by cross-validating results using both VADER [30] and TextBlob [31]. The use of open-source tools also reduced the need for resource-intensive retraining, enhancing sustainability. Efforts were made to contextualise findings carefully, avoiding overgeneralisations about cultural traits or sentiments. Interpretations were framed to respect regional diversity and avoid judgement. Transparency was prioritised throughout the research process, with methodological steps and results shared openly, fostering reproducibility and building trust in the study's outcomes.

5 Conclusion

Several key insights have emerged with actionable implications for various stakeholders within the research framework. Emotional tone in songs varies significantly by country, highlighting universal themes and cultural interpretations: while positive sentiments are globally prominent, their intensity differs regionally, suggesting music reflects and influences collective sentiment. This insight presents an opportunity for the music industry to create culturally tailored content, like localized playlists and region-specific song recommendations, to better connect with audiences. Secondly, the relationship between lyrical sentiments and happiness indices is nuanced, influenced by the sensitivity of sentiment analysis tools. Although positive sentiments typically correspond to higher happiness scores, countries with mostly neutral sentiments complicate this link. This suggests the need for refining sentiment analysis tools to better understand culturally specific emotional expressions, especially by incorporating linguistic and cultural adjustments for non-Western frameworks. Thirdly, the hypothesis that songs originating from nations experiencing turmoil are indicative of negative sentiments was not entirely supported; instead, data indicates that economically and politically challenged countries often express resilience and positivity in their lyrics as a coping mechanism. This trend provides insights for music platforms looking to improve user engagement by curating culturally aligned music suggestions and mood-driven playlists. Finally, the study highlights the complex link between public sentiment and song popularity, noting that tools often struggle to capture cultural nuances. These insights could enhance educational apps, using music to teach emotions, foster empathy, and support language acquisition. This research demonstrates that music is a multifaceted cultural artifact shaped by a confluence of emotional, linguistic, and socio-economic factors. The findings offer actionable opportunities for stakeholders across industries and also highlight the potential for further research to explore the impact of emotional tone on audience engagement and cross-cultural understanding. Future studies could benefit from integrating qualitative approaches and developing culturally adaptive sentiment models to enhance the analysis of global music trends, as well as exploring applications in user engagement, education, and beyond.

References

[1] D. Makris, “How nlp and sentiment analysis can enhance your understanding of music lyrics,” *Medium*, January 11 2023, retrieved: 23 November 2024. [Online]. Available: https://medium.com/@dimmakriss/natural-language-processing-and-sentiment-analysis-on-music-lyrics-7af53192945b

[2] D. Venditama, “Music lyrics analysis using natural language processing,” May 23 2020, retrieved: 23 November 2024. [Online]. Available: https://medium.com/analytics-vidhya/music-lyrics-analysis-using-natural-language-processing-7647922241c0

[3] M. Edwards, “The sociology of music: How music reflects and shapes social structure,” April 21 2024, retrieved: 23 November 2024. [Online]. Available: https://easysociology.com/sociology-of-media/sociology-of-music/the-sociology-of-music-how-music-reflects-and-shapes-social-structure/

[4] E. Yalcinkaya, “Music as a mirror for society: Pop song sentiment and sociopolitical factors,” September 15 2023, retrieved: 23 November 2024. [Online]. Available: https://medium.com/@eyalcinkaya/music-as-a-mirror-for-society-pop-song-sentiment-and-sociopolitical-factors-cc50e7d442b6

[5] E. Benetos, A. Ragano, D. Sgroi, and A. Tuckwell, “Measuring national happiness with music,” 2021, retrieved: 23 November 2024. [Online]. Available: https://warwick.ac.uk/fac/soc/economics/research/centres/cage/manage/publications/wp537.2021.pdf

[6] B. Li, “Research on music make people happy,” *Lecture Notes in Education Psychology and Public Media*, vol. 7, no. 1, pp. 405–410, 2023, retrieved: 23 November 2024. [Online]. Available: https://doi.org/10.54254/2753-7048/7/20221014

[7] J. Antonio, Álvaro García Martínez, P. Cano, M. Koppenberger, and F. Gouyon, “Natural language processing of lyrics,” in *ACM Multimedia*, 2005, retrieved: 23 November 2024. [Online]. Available: https://doi.org/10.1145/1101149.1101255

[8] “Spotify,” 2024, retrieved: 23 November 2024. [Online]. Available: https://open.spotify.com/genre/0JQ5DAudkNjCgYMM0TZXDw

[9] “Nltk documentation,” 2024, retrieved: 23 November 2024. [Online]. Available: https://www.nltk.org/

[10] “World happiness report 2024,” 2024, world Happiness Report. Retrieved: 21 December 2024. [Online]. Available: https://worldhappiness.report/

[11] W. F. Thompson, “Cross-cultural similarities and differences,” retrieved: 23 November 2024. [Online]. Available: https://philarchive.org/archive/BALCSA

[12] T. Fritz, S. Jentschke, N. Gosselin, D. Sammler, and I. Peretz, “Universal recognition of three basic emotions in music,” *Current Biology*, vol. 19, no. 7, pp. 573–576, 2009. [Online]. Available: https://doi.org/10.1016/j.cub.2009.02.058

[13] P. N. Juslin, G. T. Barradas, M. Ovsiannikow, J. Limmo, and W. F. Thompson, “Prevalence of emotions, mechanisms, and motives in music listening: A comparison of individualist and collectivist cultures,” *Psychology of Music*, vol. 44, no. 4, pp. 897–915, 2016. [Online]. Available: https://doi.org/10.1177/0305735615590449

[14] E. Y. Frizzell and L. C. Windsor, “Effects of teaching experience and culture on choral directors’ descriptions of choral tone,” *PLOS ONE*, vol. 16, no. 8, p. e0256587, 2021. [Online]. Available: https://doi.org/10.1371/journal.pone.0256587

[15] B. Mesquita and N. H. Frijda, “Cultural variations in emotions: A review,” *Psychological Bulletin*, vol. 112, no. 2, pp. 179–204, 1992. [Online]. Available: https://doi.org/10.1037/0033-2909.112.2.179

[16] Y. Qi and Z. Shabrina, "Sentiment analysis using twitter data: a comparative application of lexicon- and machine-learning-based approach," *Soc. Netw. Anal. Min.*, vol. 13, p. 31, 2023. [Online]. Available: https://doi.org/10.1007/s13278-023-01030-x

[17] M. Ungar, "Resilience across cultures," *British Journal of Social Work*, vol. 38, pp. 218–235, 02 2008. doi: 10.1093/bjsw/bc1343

[18] M. Eid and E. Diener, *Norms for Experiencing Emotions in Different Cultures: Inter- and Intranational Differences*, 06 2009, vol. 81, pp. 169–202. ISBN 978-90-481-2351-3

[19] D. Matsumoto and H. Hwang, "Culture and emotion," *Journal of Cross-cultural Psychology - J CROSS-CULT PSYCHOL*, vol. 43, pp. 91–118, 01 2012. doi: 10.1177/0022022111420147

[20] M. Pickering, "Taking popular music seriously. by simon frith. aldershot: Ashgate, 2007. 343 pp. isbn 0754626792," *Popular Music*, vol. 29, pp. 473 – 475, 10 2010. doi: 10.1017/S0261143010000267

[21] K. Negus and P. Astor, "Songwriters and song lyrics: Architecture, ambiguity and repetition," *Popular Music*, vol. 34, pp. 226–244, 05 2015. doi: 10.1017/S0261143015000021

[22] B. Mesquita, J. De Leersnyder, and M. Boiger, *The cultural psychology of emotion*, 01 2016, pp. 393–411.

[23] L. Sagiv, S. Roccas, J. Cieciuch, and S. Schwartz, "Personal values in human life," vol. 1, 08 2017. doi: 10.1038/s41562-017-0185-3

[24] A. Pavlenko, "Emotion and emotion-laden words in the lexicon," *Bilingualism: Language and Cognition*, vol. 11, pp. 147 – 164, 07 2008. doi: 10.1017/S1366728908003283

[25] A. Wierzbicka, "Language and metalanguage: Key issues in emotion research," *Emotion Review*, vol. 1, pp. 3–14, 01 2009. doi: 10.1177/1754073908097175

[26] L. F. Barrett, R. Adolphs, S. Marsella, A. M. Martinez, and S. D. Pollak, "Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements," *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, vol. 20, no. 1, pp. 1–68, 2019. doi: 10.1177/1529100619832930. [Online]. Available: https://doi.org/10.1177/1529100619832930

[27] L. Besacier, E. Barnard, A. Karpov, and T. Schultz, "Automatic speech recognition for under-resourced languages: A survey," *Speech Communication*, vol. 56, p. 85–100, 01 2014. doi: 10.1016/j.specom.2013.07.008

[28] "End user license agreement (eula)," 2021, musixmatch. Retrieved: 23 November 2024. [Online]. Available: https://about.musixmatch.com/eula

[29] "Privacy policy," 2022, genius. Retrieved: 23 November 2024. [Online]. Available: https://genius.com/static/privacy_policy

[30] "Welcome to vadersentiment's documentation," 2021, retrieved: 23 November 2024. [Online]. Available: https://vadersentiment.readthedocs.io/en/latest/

[31] "Textblob: Simplified text processing," textBlob. Retrieved: 23 November 2024. [Online]. Available: https://textblob.readthedocs.io/en/dev/

[32] J. F. Helliwell, R. Layard, J. D. Sachs, J.-E. D. Neve, L. B. Aknin, and S. Wang, *World Happiness Report 2024*, J. F. Helliwell, R. Layard, J. D. Sachs, J.-E. De Neve, L. B. Aknin, and S. Wang, Eds. Oxford, UK: Wellbeing Research Centre, University of Oxford, 2024. ISBN 978-1-7348080-7-0 Retrieved: 21 December 2024.

[33] A. Singh, "World happiness report 2024 dataset," 2024, retrieved: 20 November 2024 from Kaggle. [Online]. Available: https://www.kaggle.com/datasets/ajaypalsinghlo/world-happiness-report-2024

A List of available languages with the NLTK package for sentiment analysis (NLTK v.3.4.5)

```
['arabic',
 'azerbaijani',
 'danish',
 'dutch',
 'english',
 'finnish',
 'french',
 'german',
 'greek',
 'hungarian',
 'indonesian',
 'italian',
 'kazakh',
 'nepali',
 'norwegian',
 'portuguese',
 'romanian',
 'russian',
 'slovene',
 'spanish',
 'swedish',
 'tajik',
 'turkish']
```

Figure 4: Terminal result of the call to the function `print(stopwords.fileids())` which returns all the available language packages in NLTK

B Structure of sentiment analysis scores of TextBlob and VADER tools

TextBlob Results

TextBlob scores were structured as follows:

- **Polarity** for sentiment, it's a float that ranges between -1 (indicating negative sentiment) and +1 (indicating positive sentiment).
- **Subjectivity** refers to personal opinion expressed in a text content and it's quantified using a float that lies in the range (0,1). If the value of subjectivity is above 0.5, the sentence is more subjective whereas when the value is below 0.5 is more objective.

VADER Results

VADER scores were structured as follows

- **Neutral Score** for neutral sentiment, the value goes from 0 to 1.
- **Positive Score** for positive sentiment, the value goes from 0 to 1;
- **Negative Score** for negative sentiment, the value goes from 0 to 1
- **Compound Score** for an overall score that combines negative, positive, and neutral sentiments into a single score. This was calculated using the sum of all normalised ratings between -1 and +1 for most negative and most positive respectively

C Structure of happiness rankings from World Happiness Report

The World Happiness Report is a global survey assessing happiness levels across countries [10]. This report has gained global recognition, with policymakers increasingly using happiness metrics to guide decisions. The report highlights how the science of happiness explains differences in individual and national happiness. Happiness scores and rankings are based on Gallup World Poll data, which includes responses to the ”Cantril ladder” question. This asks participants to rate their current life on a scale from 0 (worst possible) to 10 (best possible). Scores are derived from nationally representative samples and adjusted using Gallup's weighting system. Key factors contributing to happiness include [32]:

- Social Support
- Life Expectancy
- Freedom
- Generosity
- Absence of Corruption
- Economic Production (GDP per capita)

These factors are compared against a hypothetical benchmark called Dystopia, representing the lowest global averages for all six metrics. While these factors explain variations in rankings, they do not directly affect the total happiness score. Dystopia is a conceptual country with the least-happy conditions: lowest income, life expectancy, generosity, freedom, and highest corruption. It provides a baseline for comparison, ensuring all country scores are positive. The Dystopia Residual represents unexplained differences between observed happiness and predictions based on the six factors.

It was used the ”World Happiness Report 2024” dataset, retrieved from Kaggle [33]

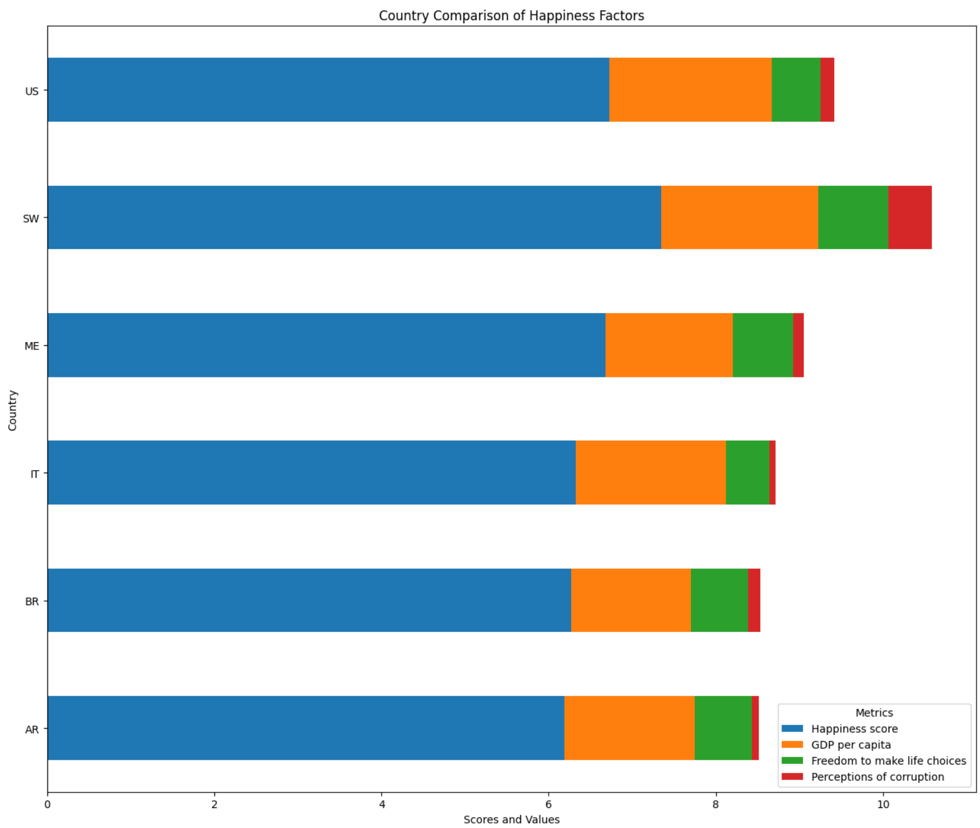

Figure 5: Comparison of Happiness Factors and Sub-Indices per Country

D VADER Analysis

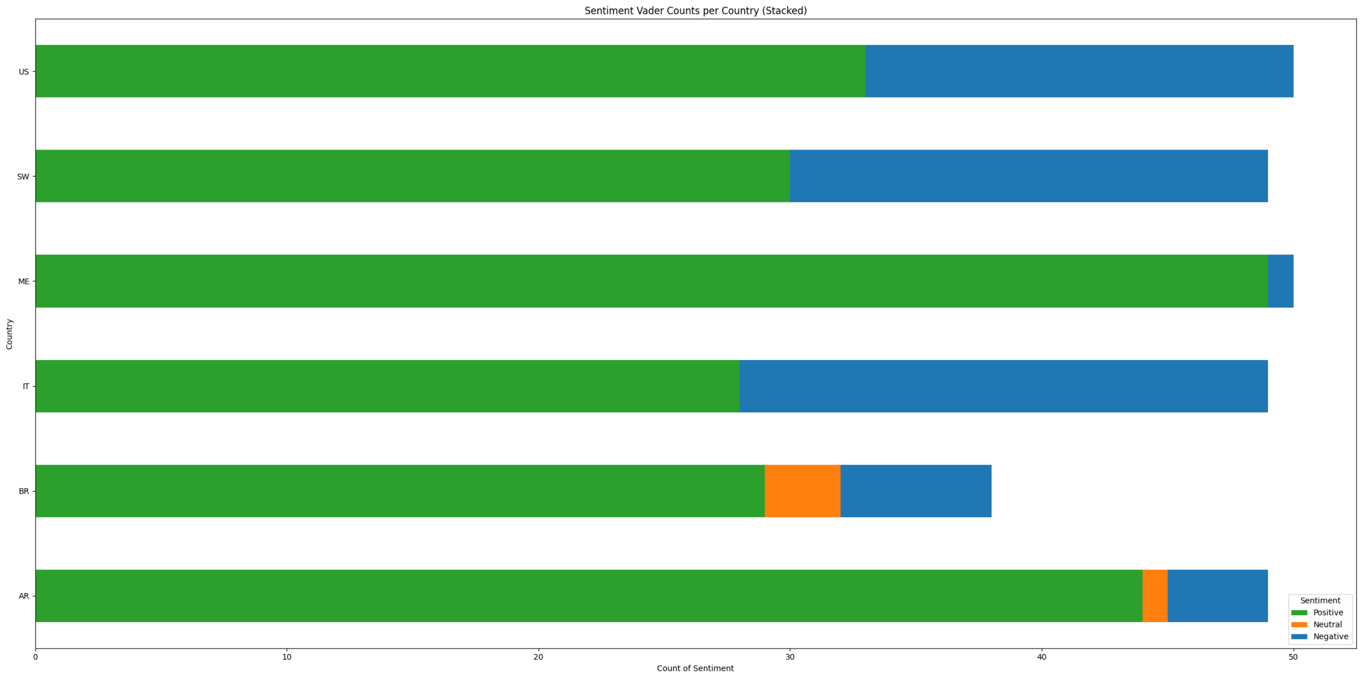

Figure 6: VADER sentiment distribution (positive, neutral, negative) per country

Country	Total Sum	Mean	Median	Std dev	Min Value	Max Value
AR	30.6505	0.625520	0.75790	0.447431	-0.824463	0.9999
BR	14.3205	0.376855	0.36120	0.447431	-0.9922	0.9999
IT	0.9891	0.020186	0.10270	0.641737	-0.9982	0.9984
ME	28.2604	0.565208	0.49670	0.309678	-0.2732	0.9911
SW	9.4864	0.193600	0.57190	0.824463	-0.9888	0.9993
US	17.1159	0.342318	0.57190	0.824463	-0.9989	0.9993

Table 2: Summary Statistics for VADER Compound Score by Country

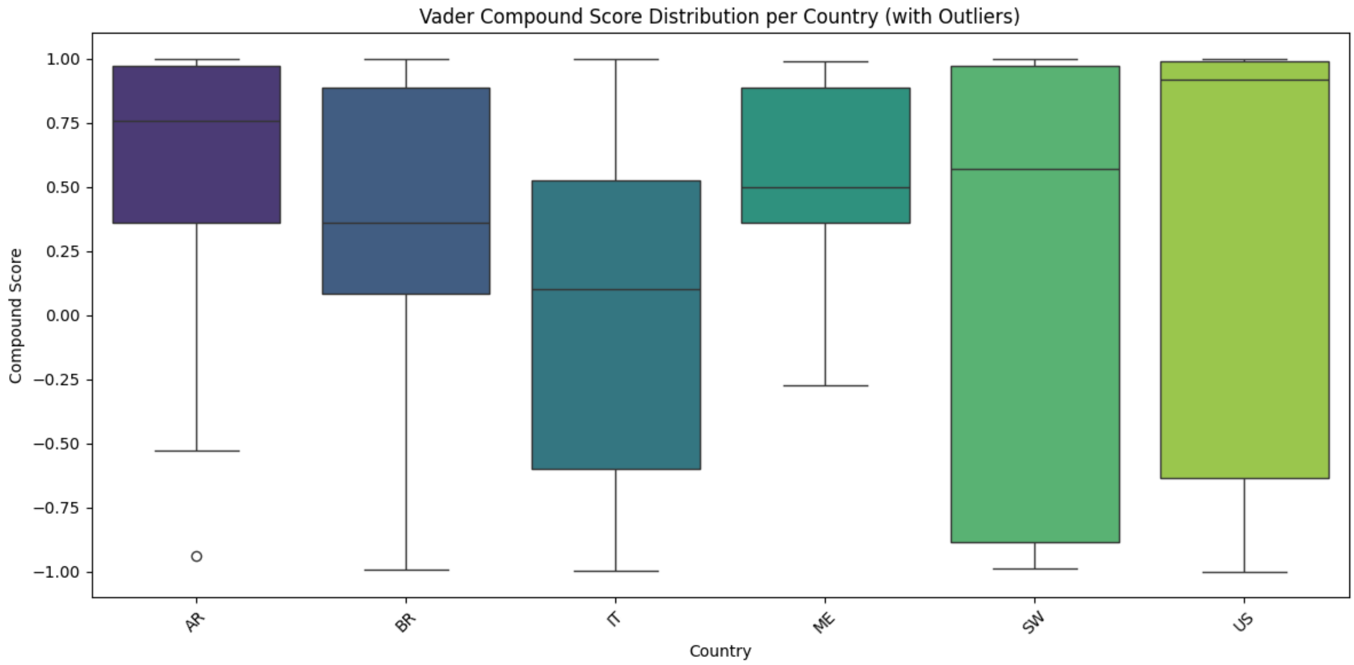

Figure 7: VADER compound score distribution per country

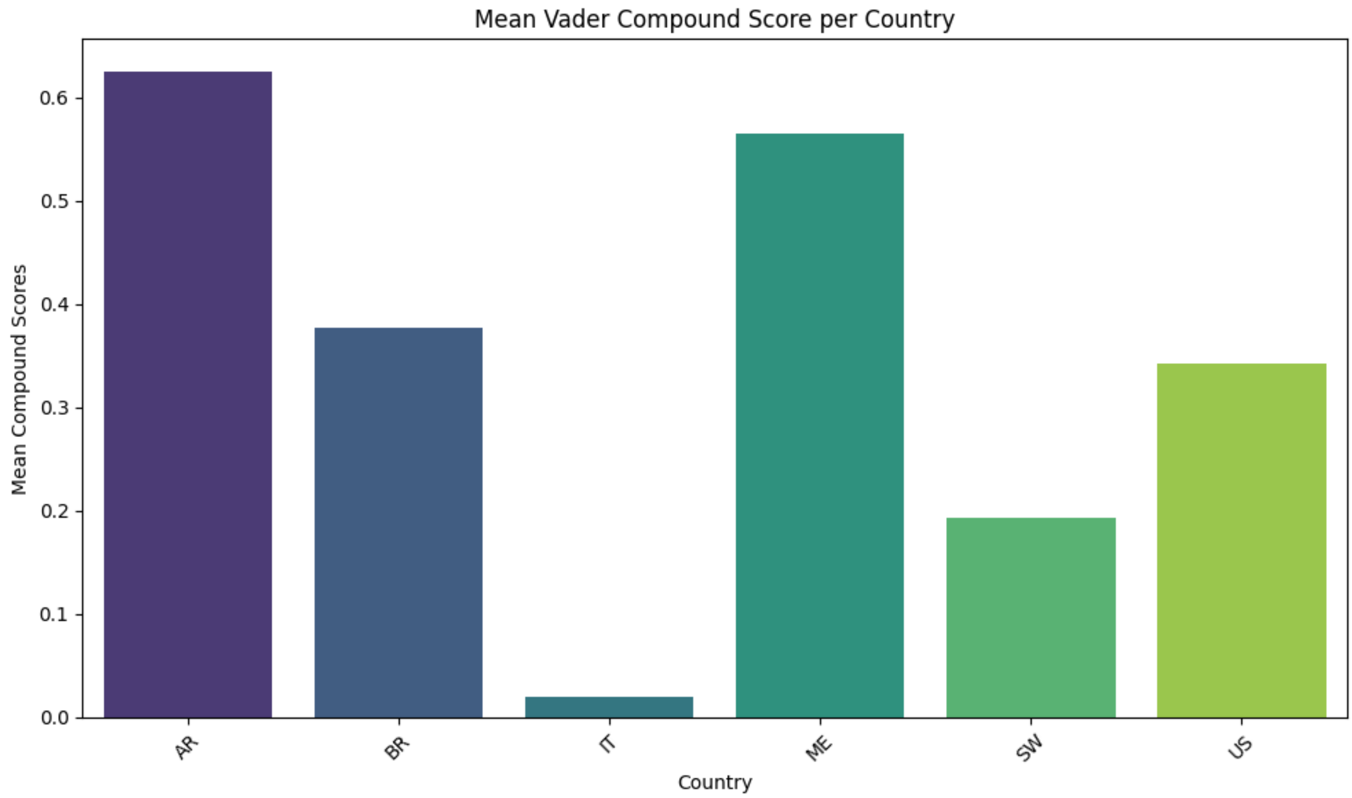

Figure 8: VADER mean compound score distribution per country

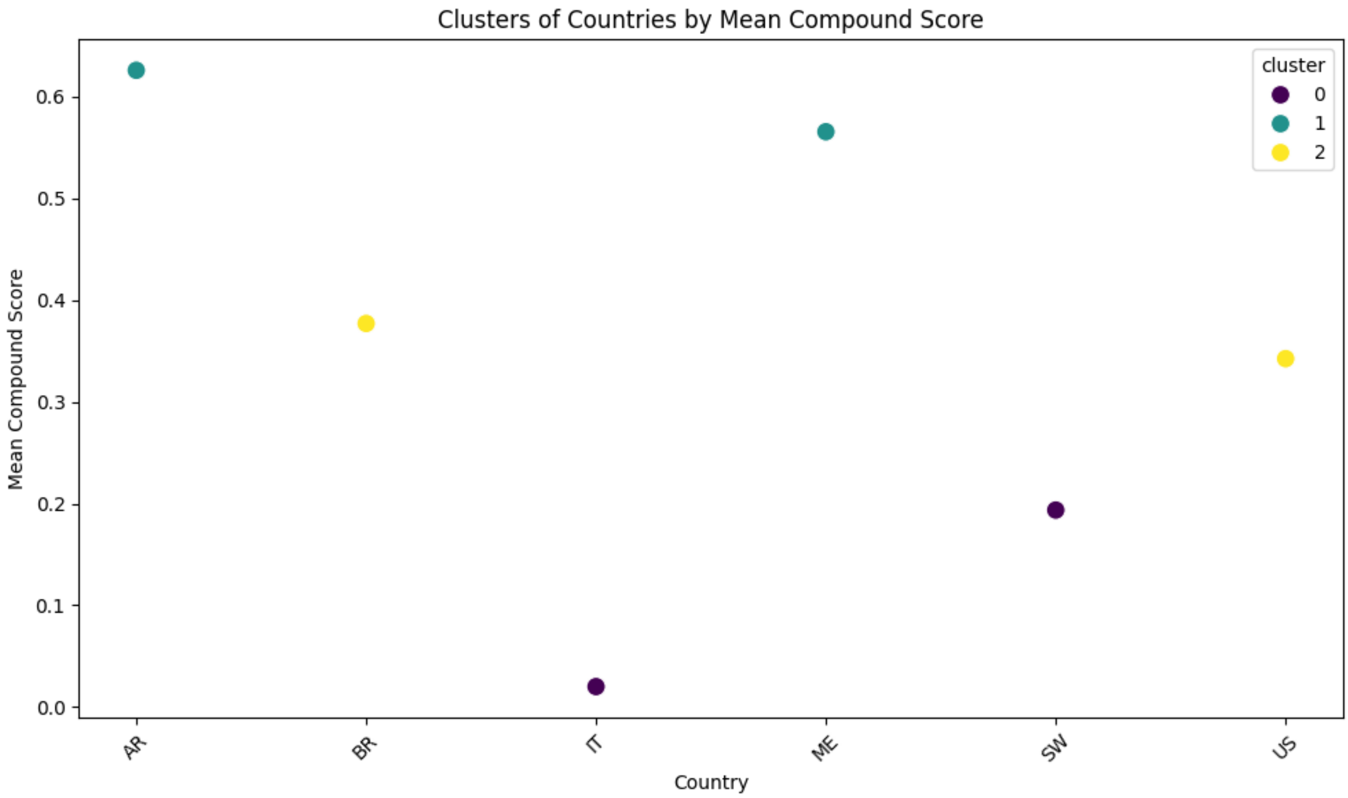

Figure 9: KMeans Clustering by countries based on VADER mean compound score

E TextBlob Analysis

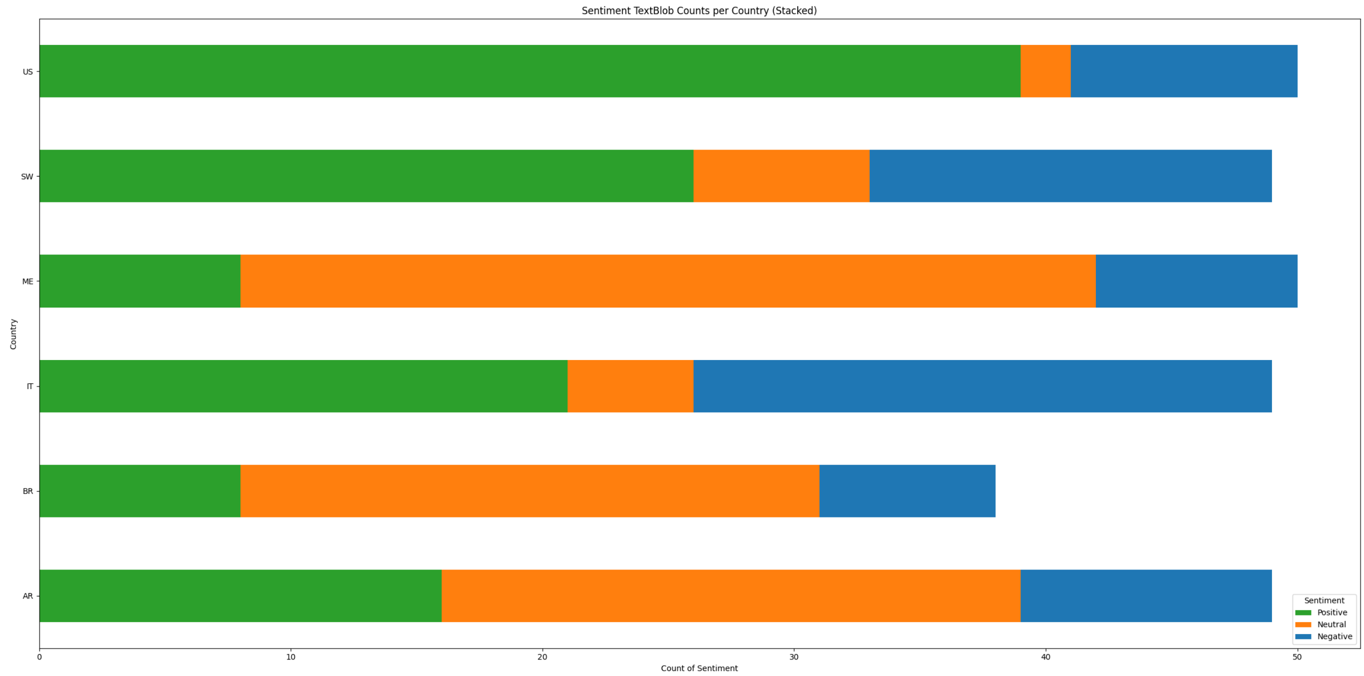

Figure 10: TextBlob sentiment counts per country (positive, neutral, negative)

Country	Total Sum	Mean	Median	Std dev	Min Value	Max Value
AR	0.7218	0.014731	0	0.196631	-0.45	0.428571
BR	0.45517	0.011978	0	0.123268	-0.244444	0.5
IT	-0.422823	-0.008629	0	0.172869	-0.444242	0.396639
ME	0.000302	0.000006	0	0.119445	-0.5	0.35
SW	2.006452	0.040948	0.035714	0.221342	-0.6	0.6
US	5.072849	0.101457	0.068032	0.17587	-0.5	0.458333

Table 3: Summary Statistics for TextBlob Polarity Score by Country

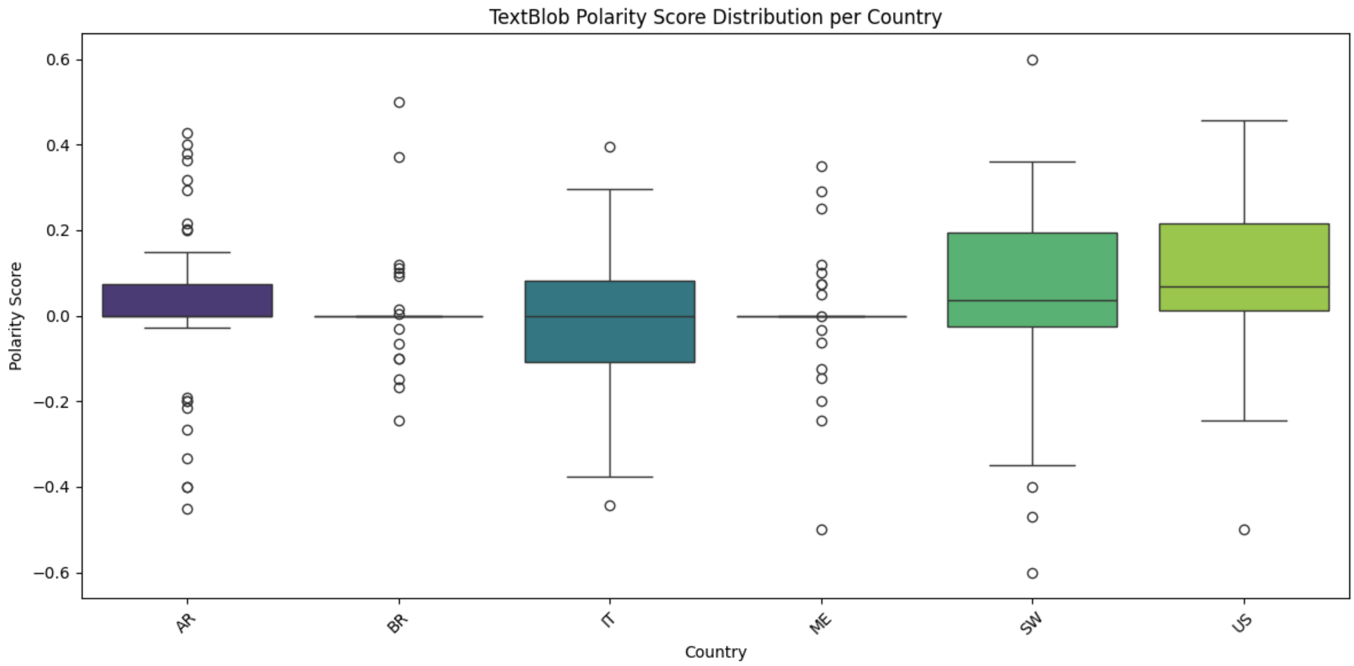

Figure 11: TextBlob Polarity Score distribution by Country

Country	Total Sum	Mean	Median	Std Dev	Min Value	Max Value
AR	14.358753	0.293036	0.216667	0.273353	0	0.8
BR	8.316682	0.218860	0.033333	0.289651	0	1
IT	20.6764	0.421967	0.400	0.232003	0	1
ME	10.066643	0.201333	0.061806	0.255786	0	1
SW	21.353516	0.435786	0.477778	0.242147	0	0.9
US	25.102706	0.502054	0.515696	0.170160	0	1

Table 4: Summary Statistics for TextBlob Subjectivity Score by Country

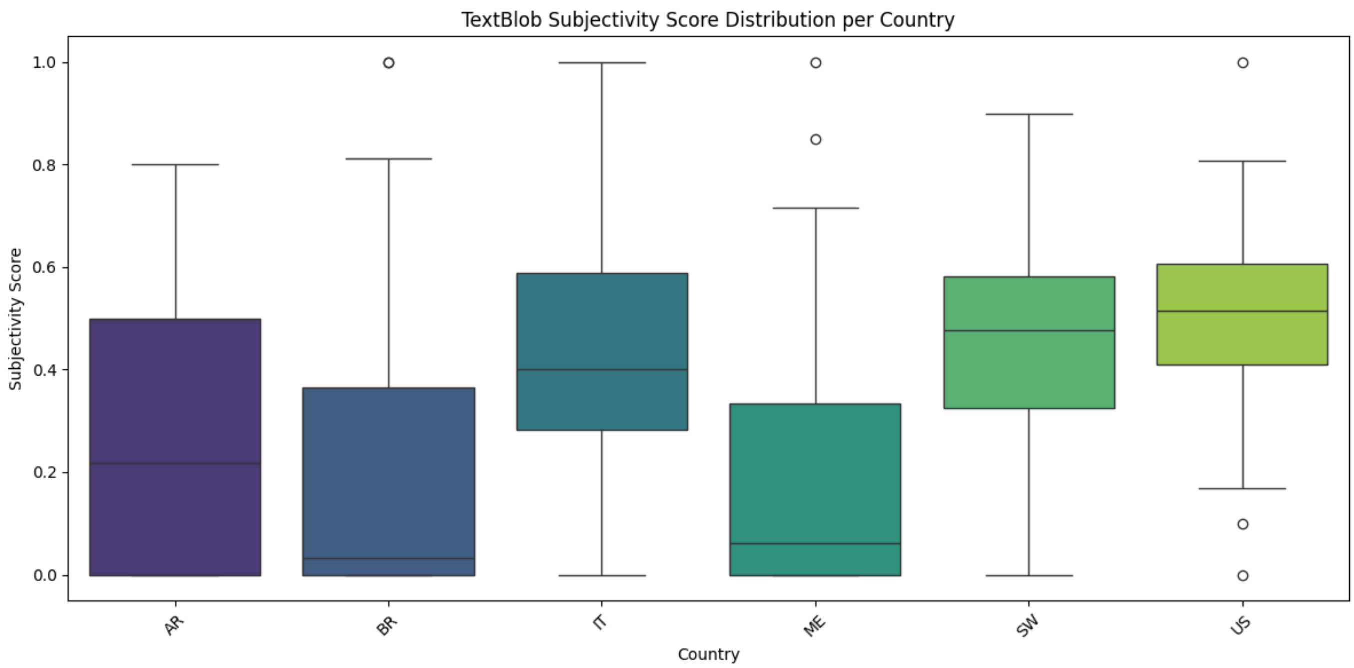

Figure 12: TextBlob Subjectivity Score Distribution by Country

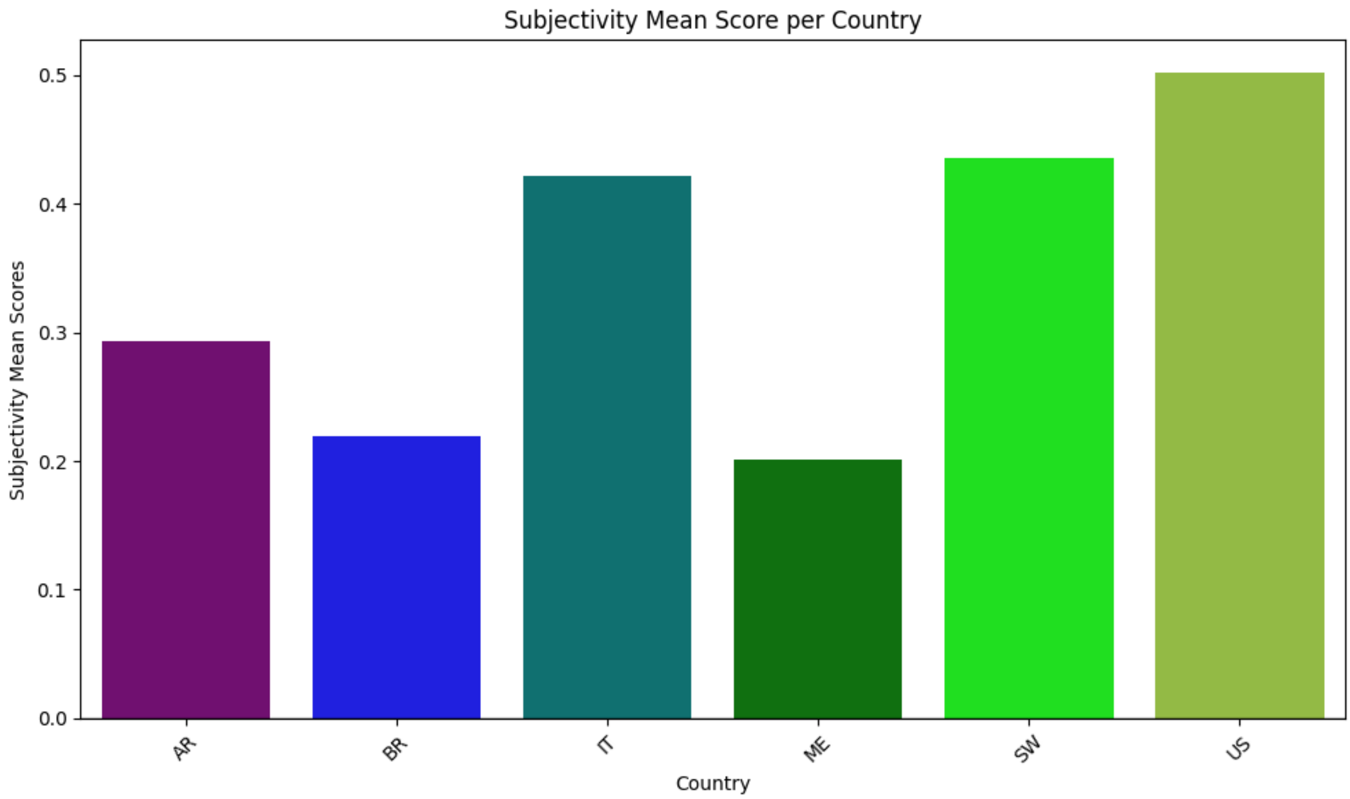

Figure 13: TextBlob mean Subjectivity Score Distribution by Country

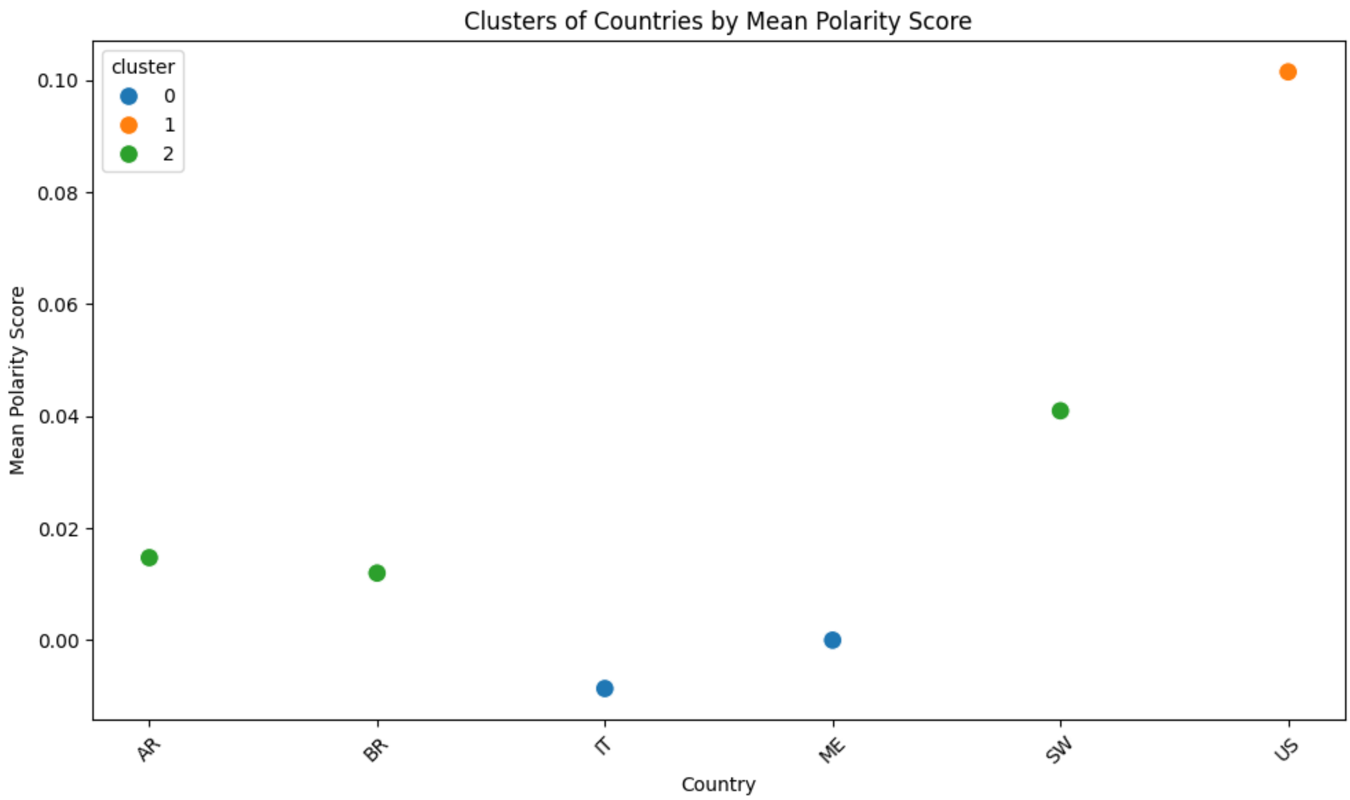

Figure 14: Clusters of Countries per TextBlob Mean Polarity Score

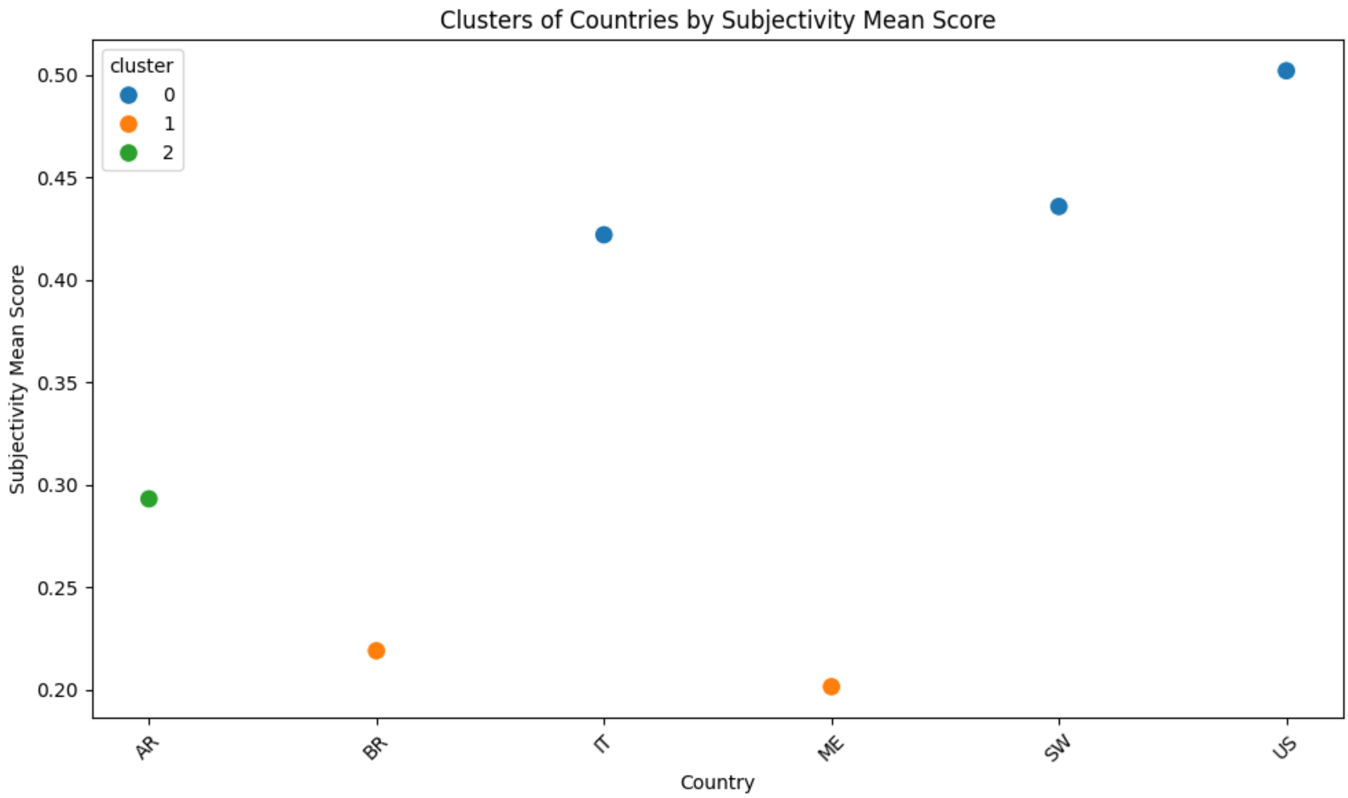

Figure 15: Clusters of Countries per TextBlob Mean Subjectivity Score

F Comparative Analysis of Happiness Score and Sentiments

Country	Happiness Score	Compound VADER	Polarity TextBlob
AR	6.188	0.625520	0.014731
BR	6.272	0.376855	0.011978
IT	6.324	0.020186	-0.008629
ME	6.678	0.565208	0.000006
SW	7.344	0.193600	0.040948
US	6.725	0.342318	0.101457

Table 5: List of Happiness Scores and Sentiment Scores by Country

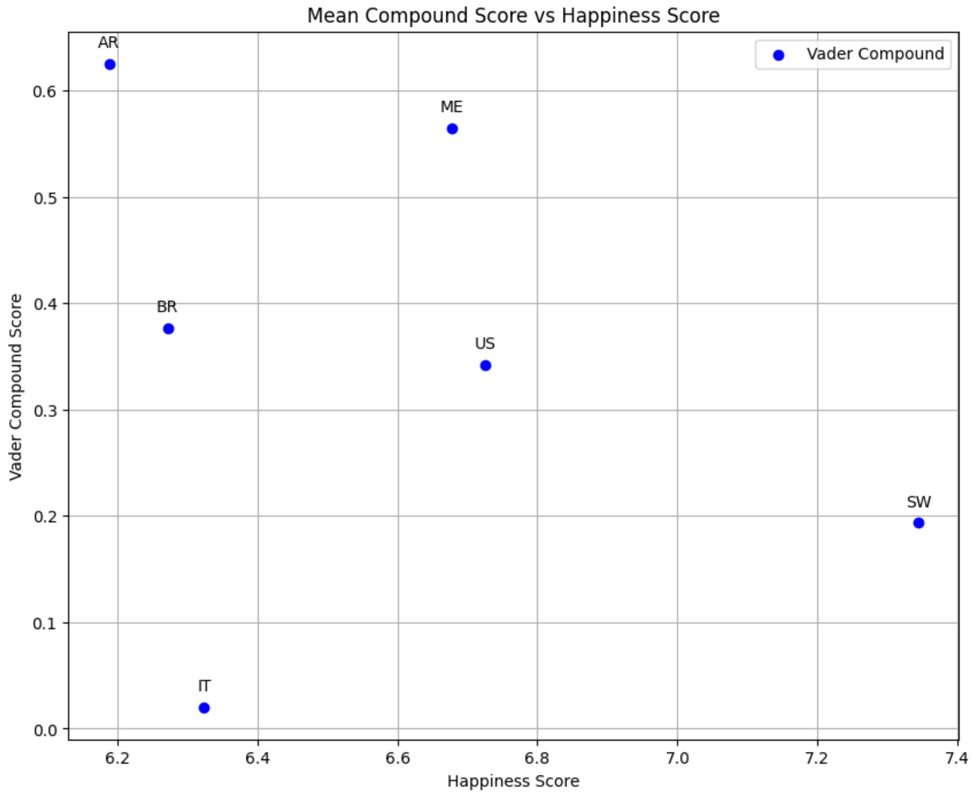

Figure 16: Scatterplot showing the relationship between VADER'S Mean Compound Score and Happiness Score per Country

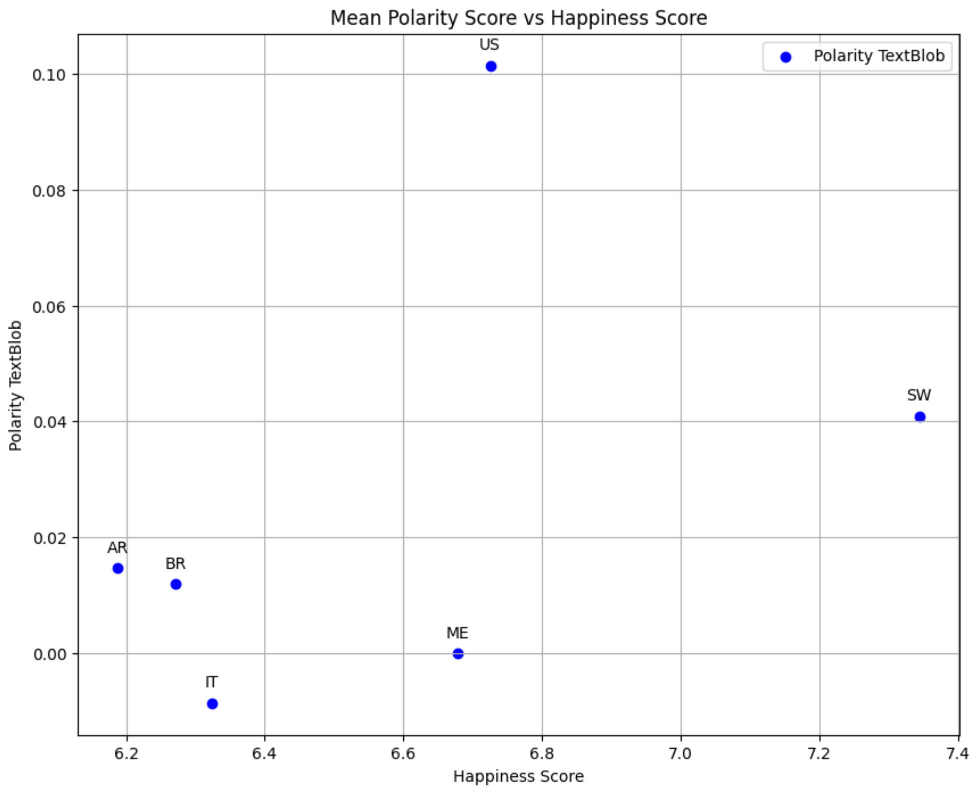

Figure 17: Scatterplot showing the relationship between TextBlob's Mean Polarity Score and Happiness Score per Country

G Validating questions and hypotheses

Country	Sentiment (VADER)	Count (VADER)	Percentage (VADER)	Sentiment (TextBlob)	Count (TextBlob)	Percentage (TextBlob)
AR	negative	4	8.16%	negative	10	20.41%
AR	neutral	1	2.04%	neutral	23	46.94%
AR	positive	44	89.80%	positive	16	32.65%
Total Count: 49						
BR	negative	6	15.79%	negative	7	18.42%
BR	neutral	3	7.89%	neutral	23	60.53%
BR	positive	29	76.32%	positive	8	21.05%
Total Count: 38						
IT	negative	21	42.86%	negative	23	46.94%
IT	neutral	0	0.00%	neutral	5	10.20%
IT	positive	28	57.14%	positive	21	42.86%
Total Count: 49						
ME	negative	1	2.00%	negative	8	16.00%
ME	neutral	0	0.00%	neutral	34	68.00%
ME	positive	49	98.00%	positive	8	16.00%
Total Count: 50						
SW	negative	19	38.78%	negative	16	32.65%
SW	neutral	0	0.00%	neutral	7	14.29%
SW	positive	30	61.22%	positive	26	53.06%
Total Count: 49						
US	negative	17	34.00%	negative	9	18.00%
US	neutral	0	0.00%	neutral	2	4.00%
US	positive	33	66.00%	positive	39	78.00%
Total Count: 50						

Table 6: VADER and TextBlob Sentiment Distribution and Comparison per Country

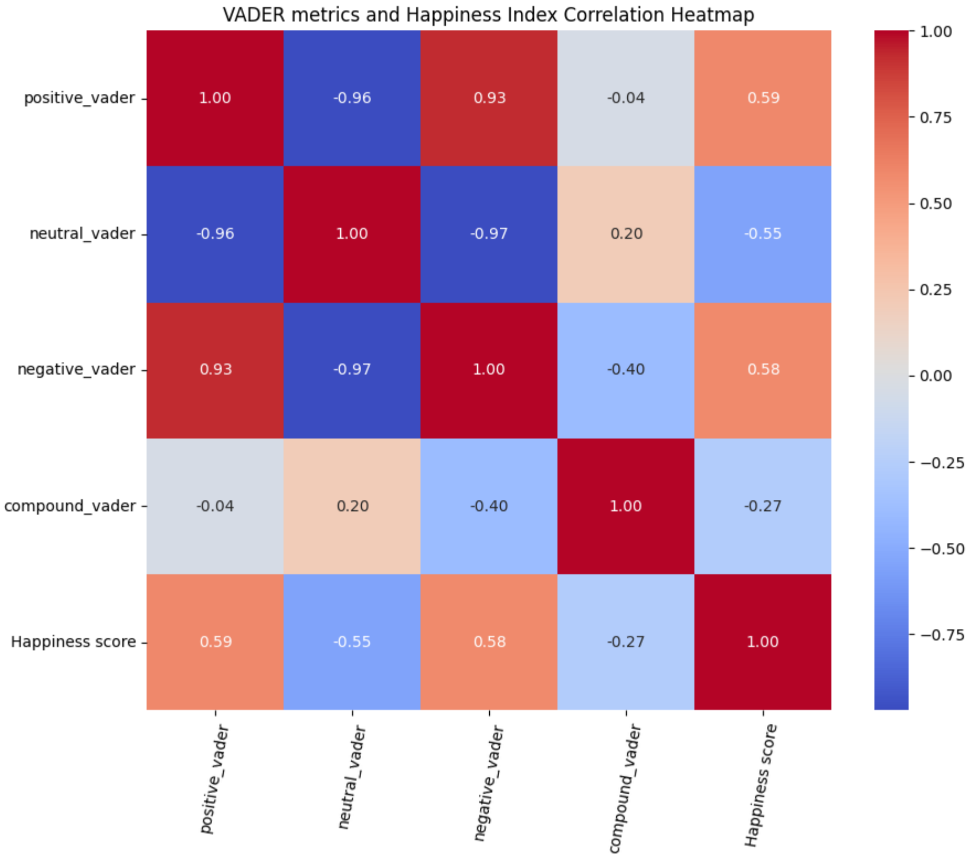

Figure 18: VADER metrics and Happiness Index Correlation Matrix

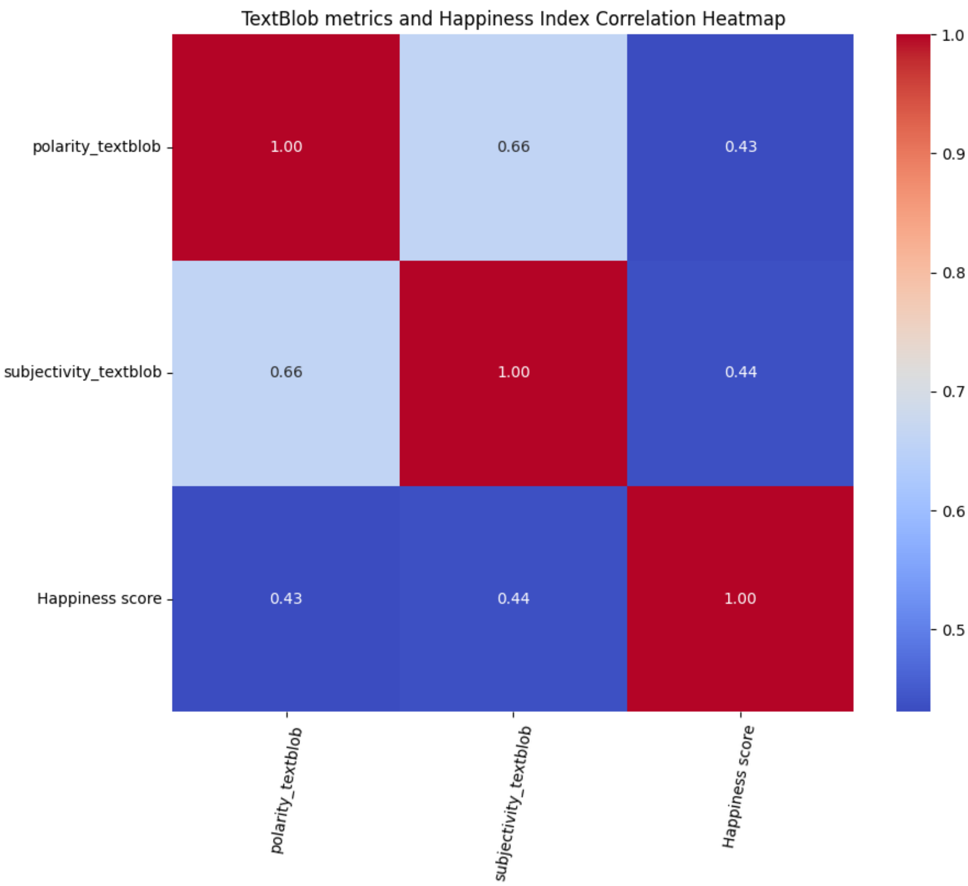

Figure 19: TextBlob metrics and Happiness Index Correlation Matrix

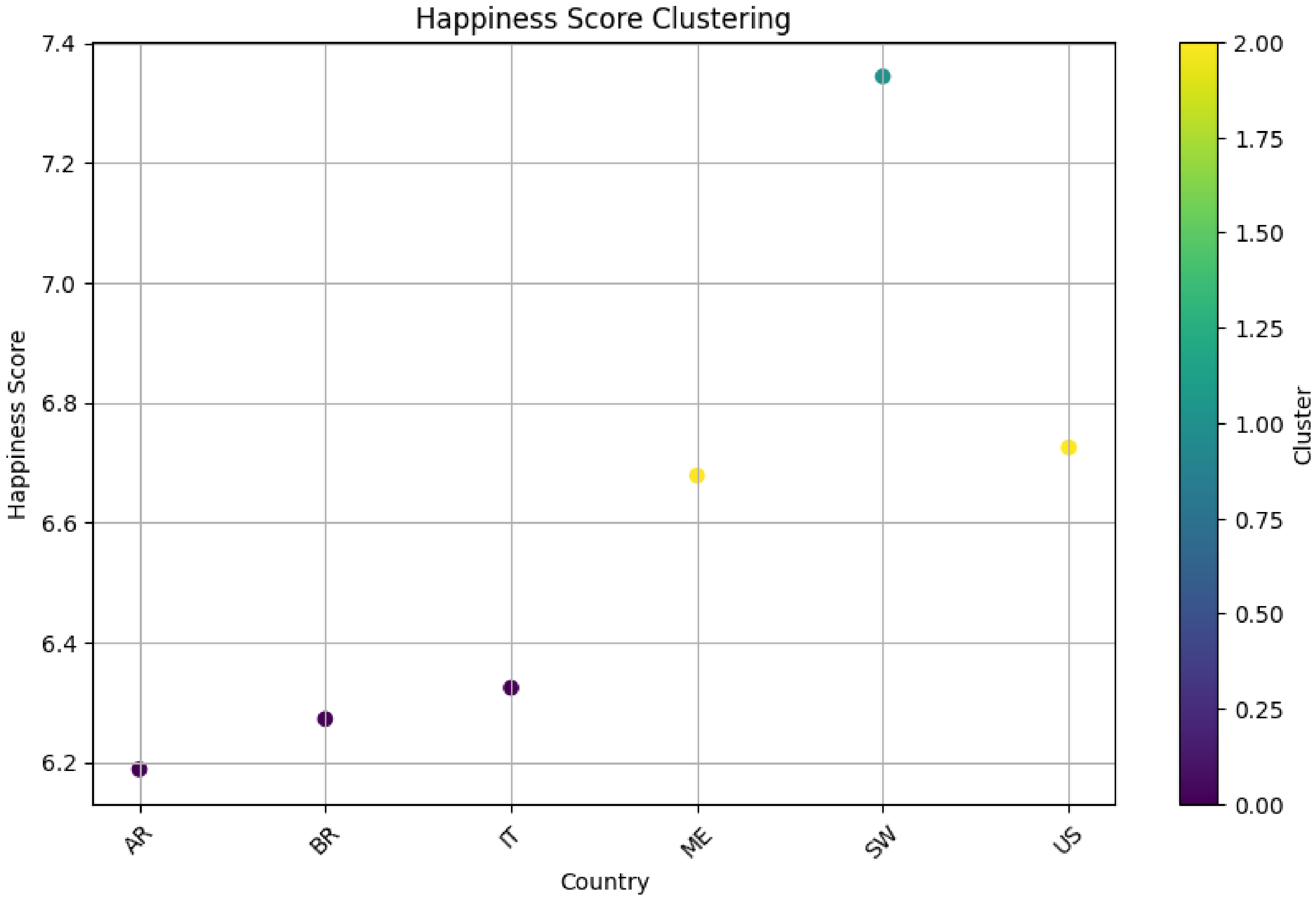

Figure 20: Clusters of Countries by Happiness Score

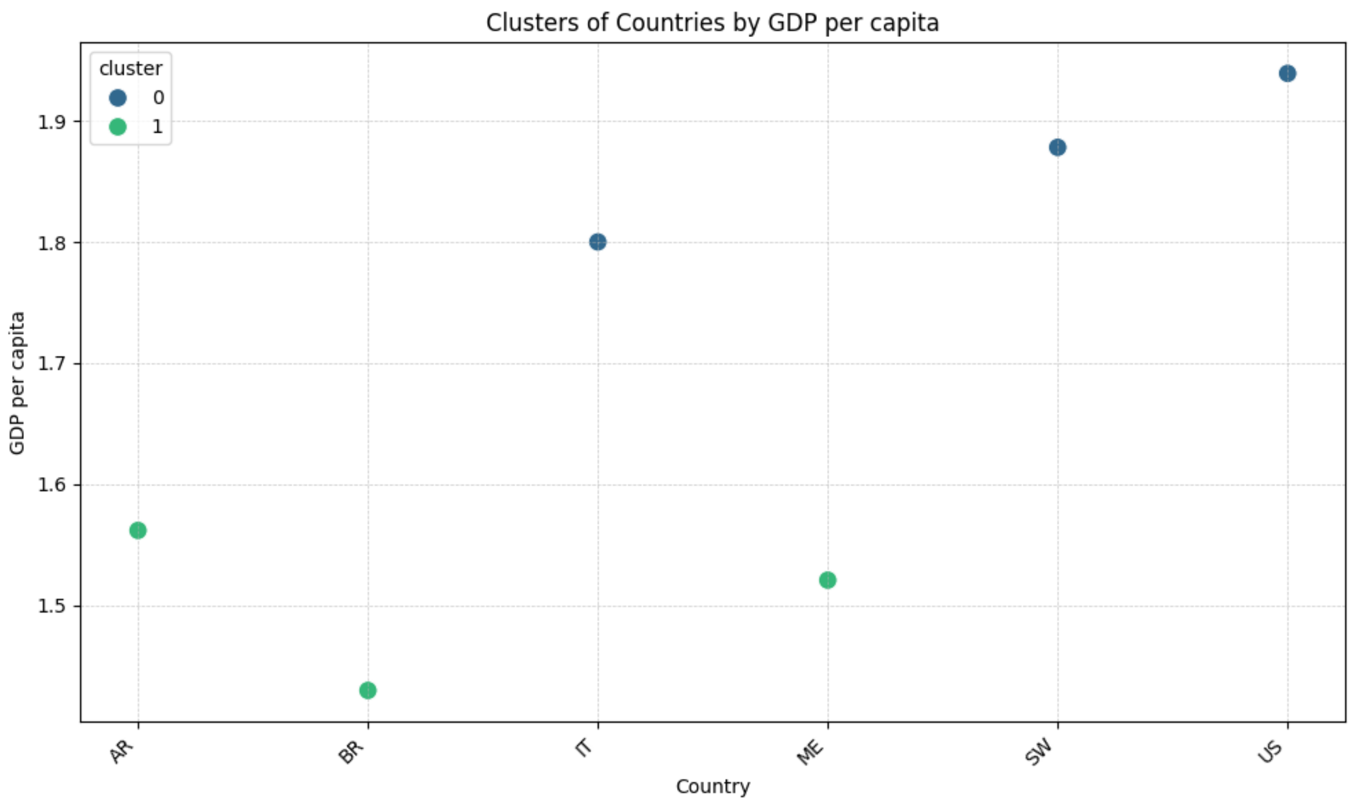

Figure 21: Clusters of Countries by GDP per capita

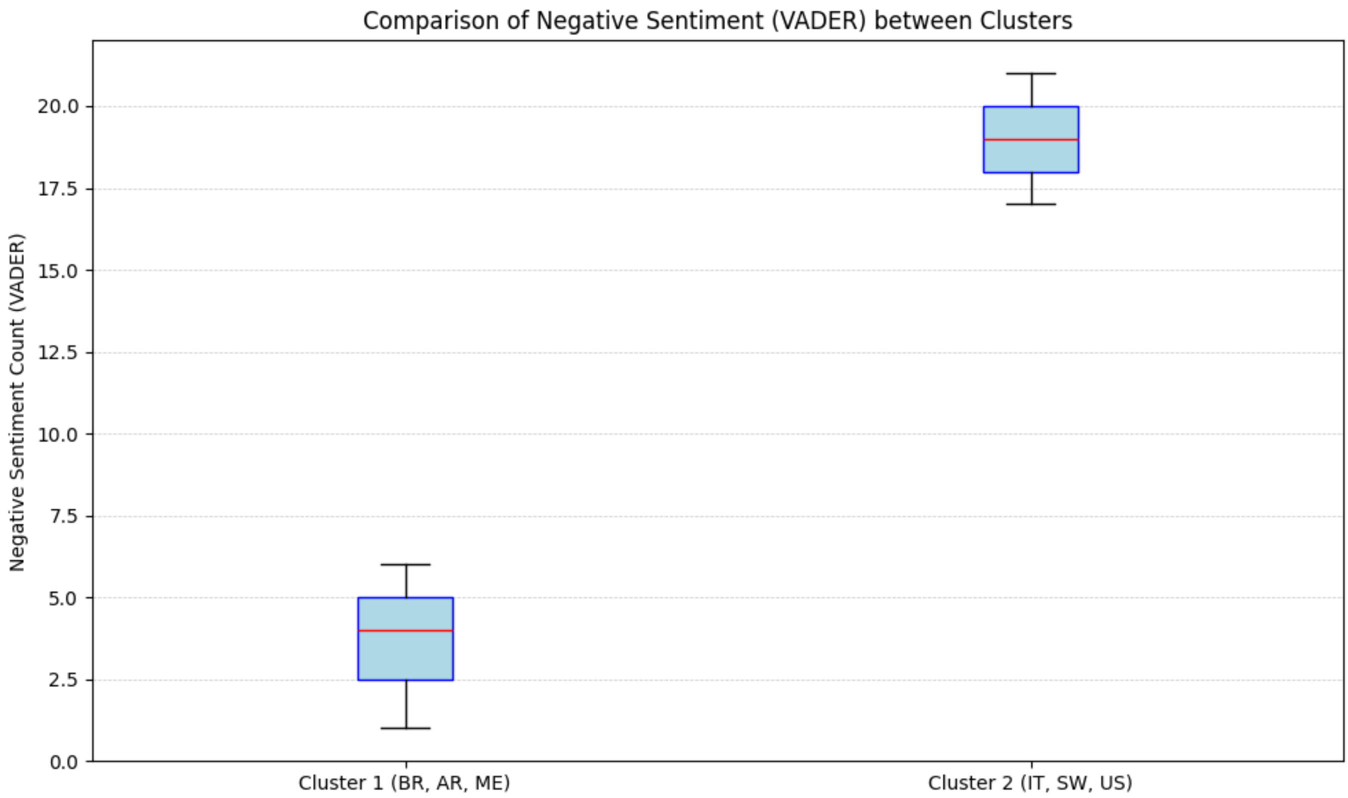

Figure 22: Comparison of VADER negative sentiment counts between economical Clusters

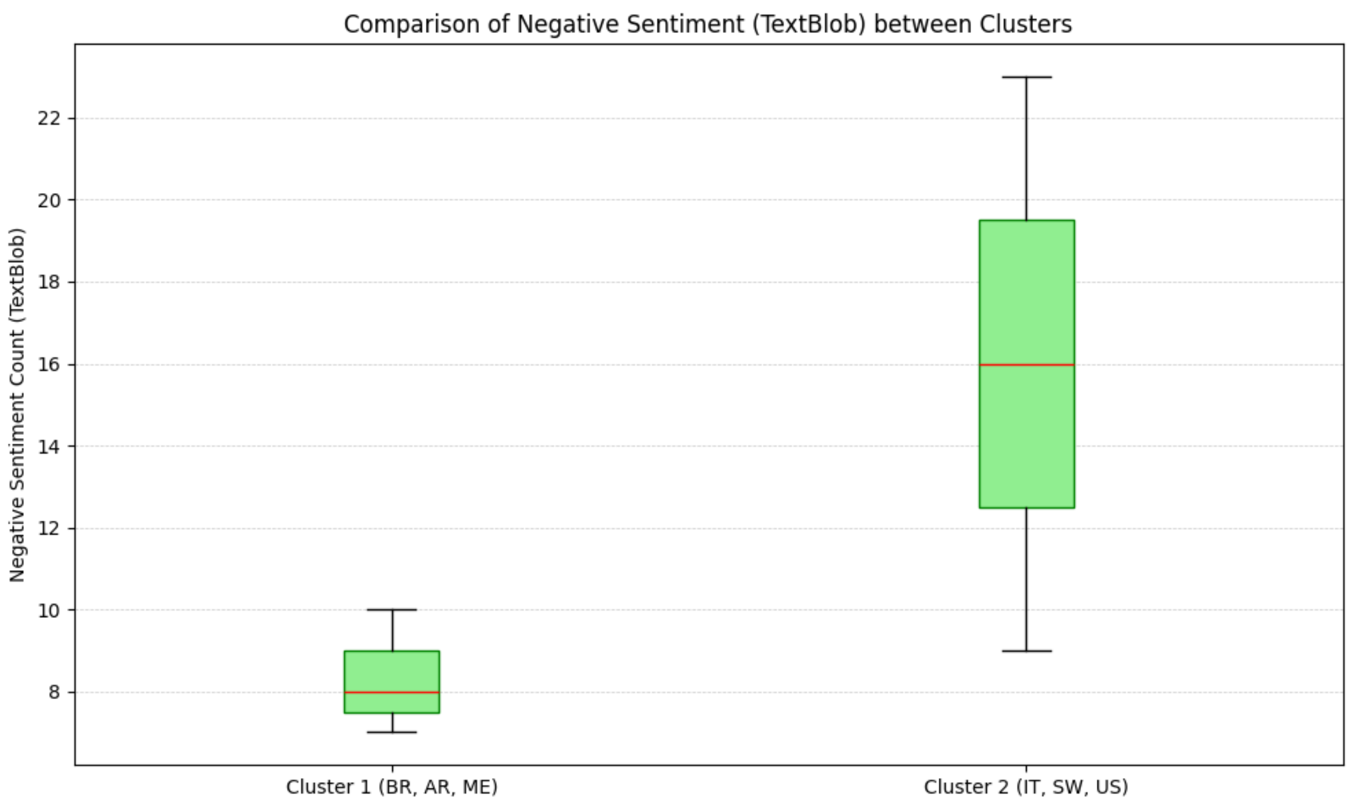

Figure 23: Comparison of TextBlob negative sentiment counts between economical Clusters

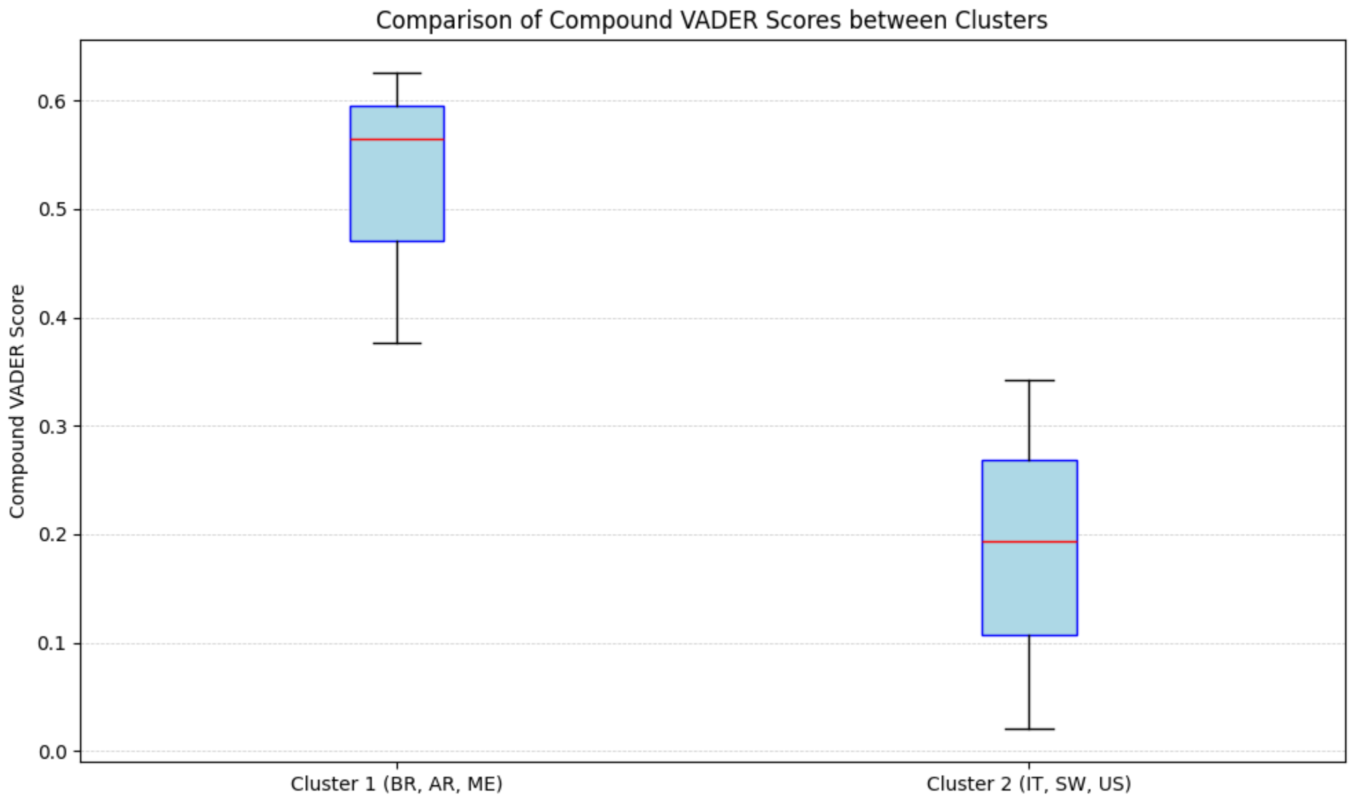

Figure 24: Comparison of VADER compound scores between economical Clusters

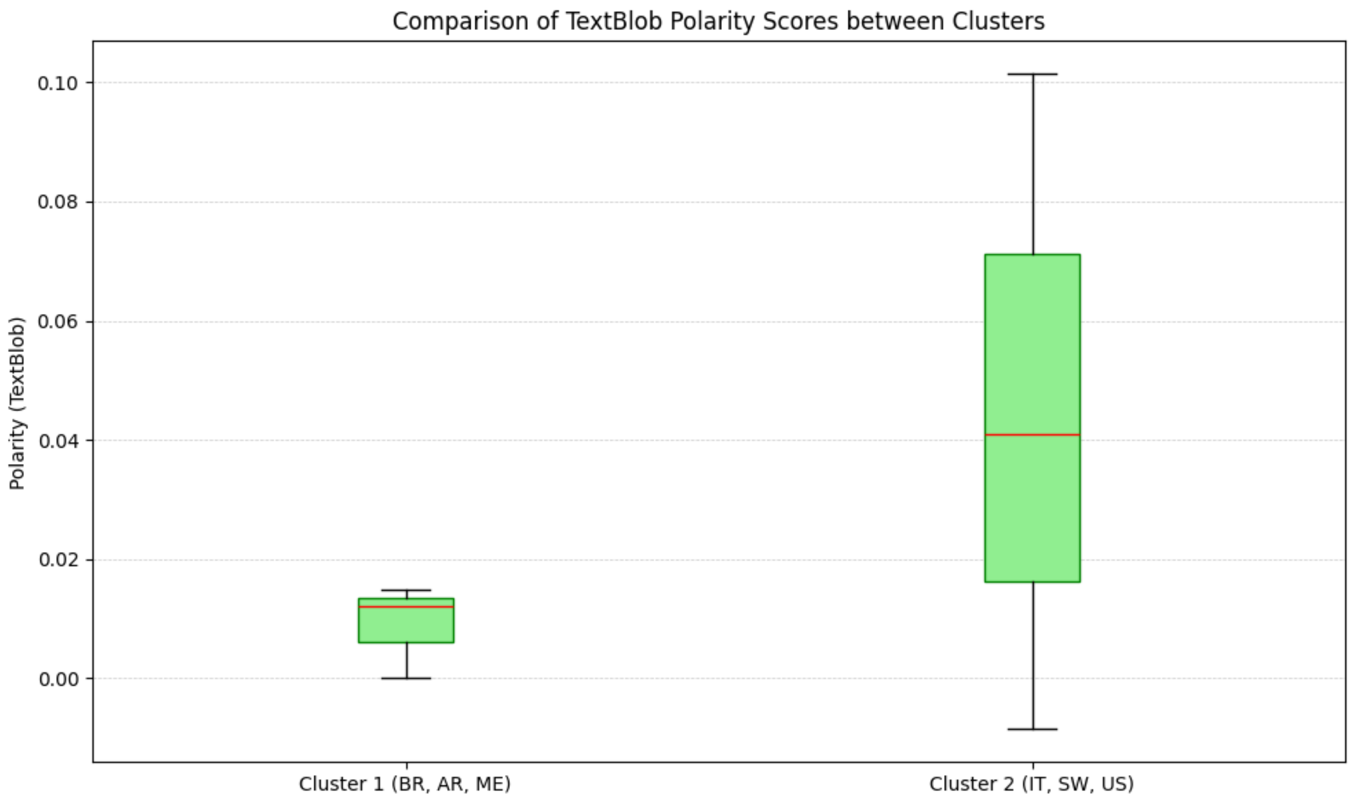

Figure 25: Comparison of TextBlob Polarity scores between economical Clusters

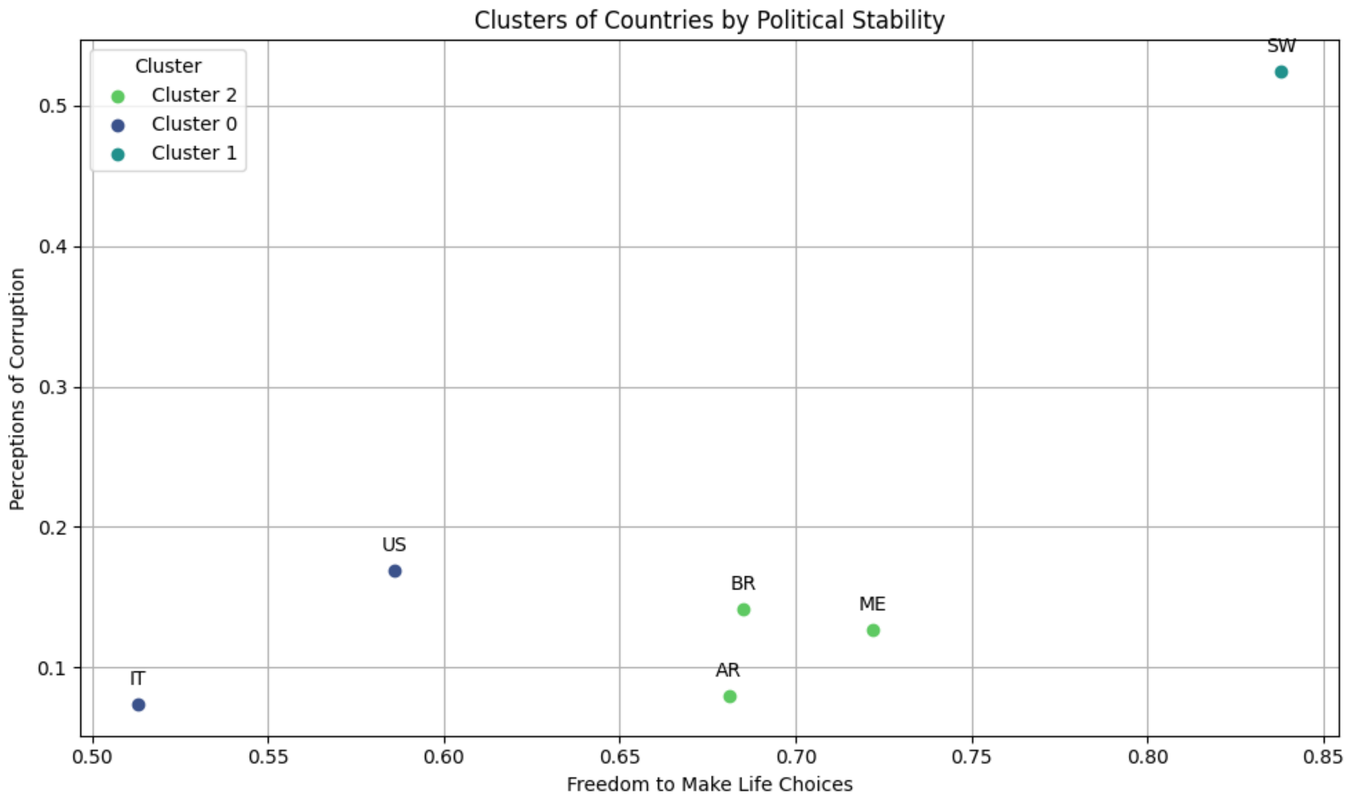

Figure 26: Clusters of Countries by Perceived Political Stability

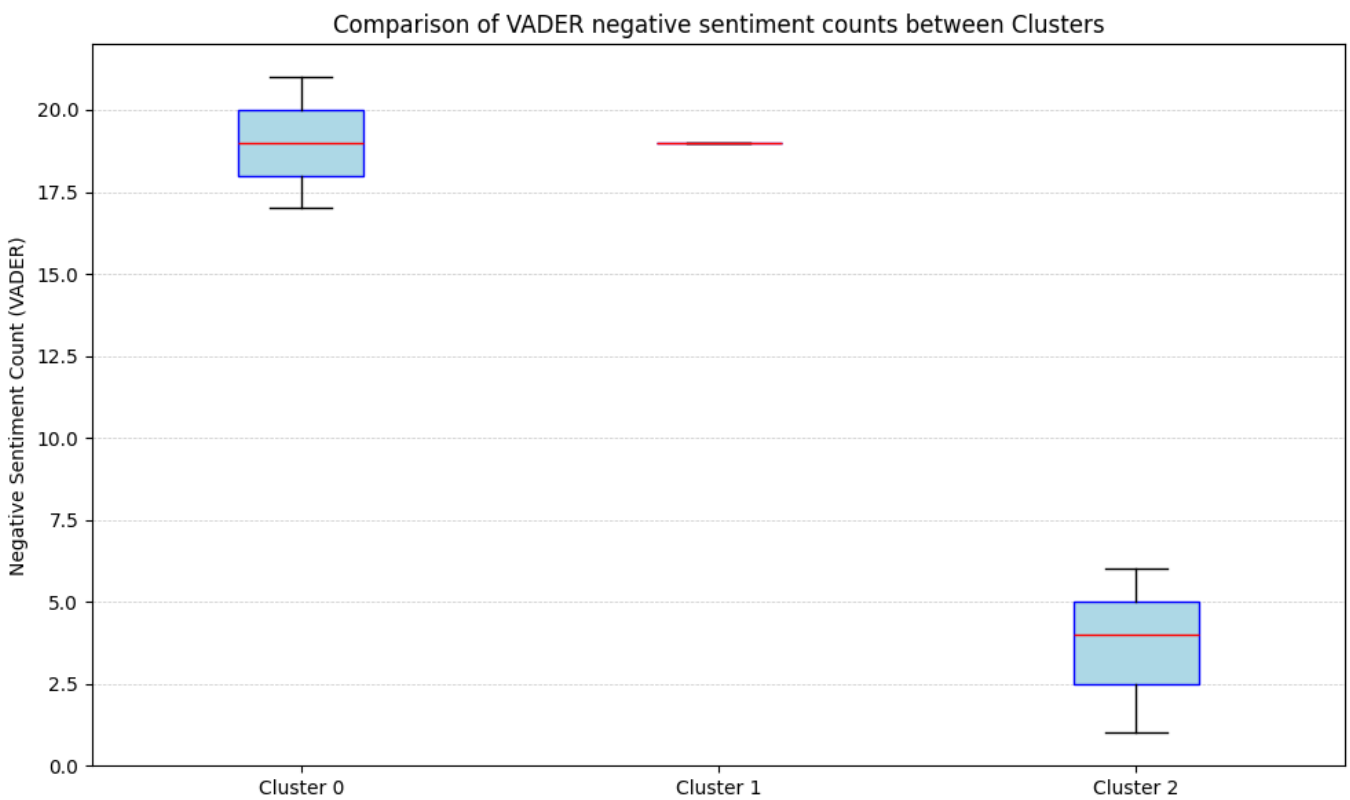

Figure 27: Comparison of VADER negative sentiment counts between political Clusters

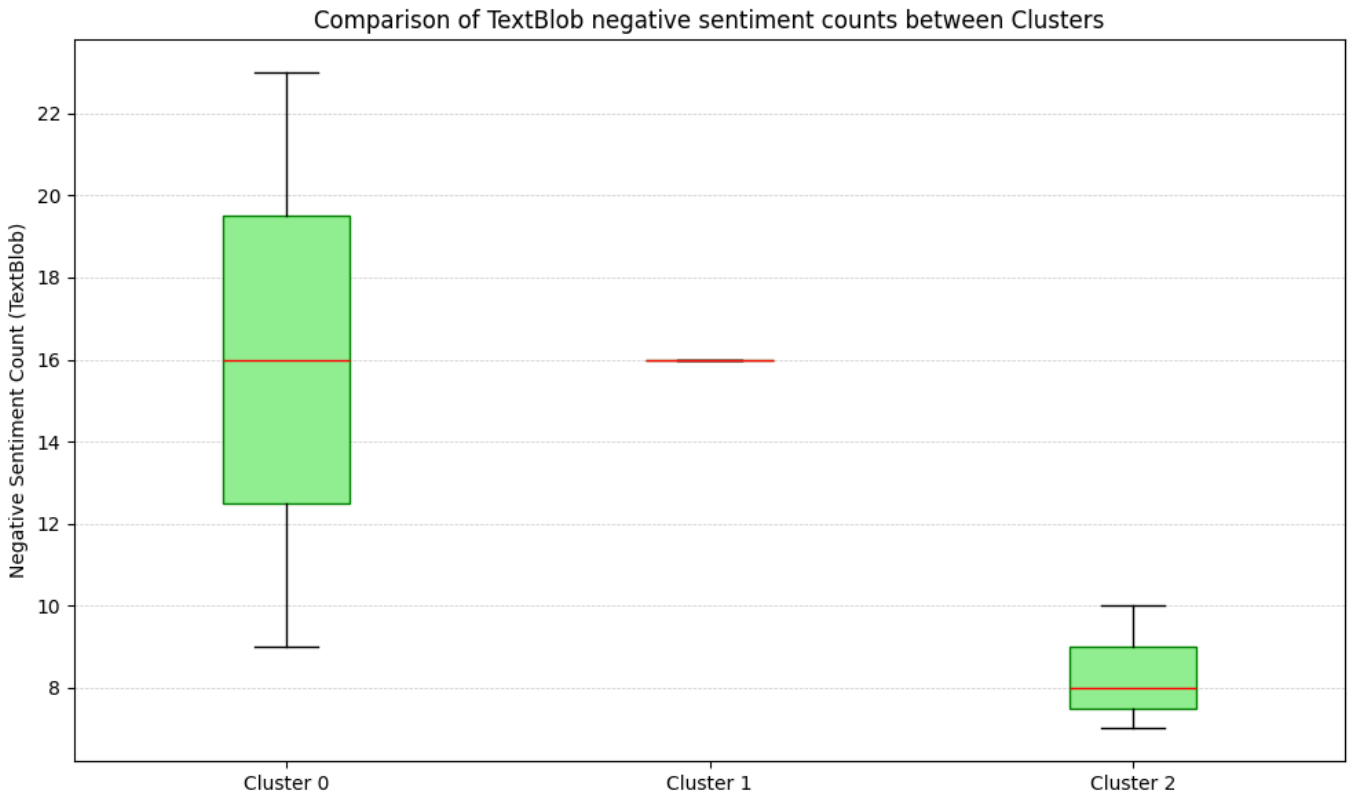

Figure 28: Comparison of TextBlob negative sentiment counts between political Clusters

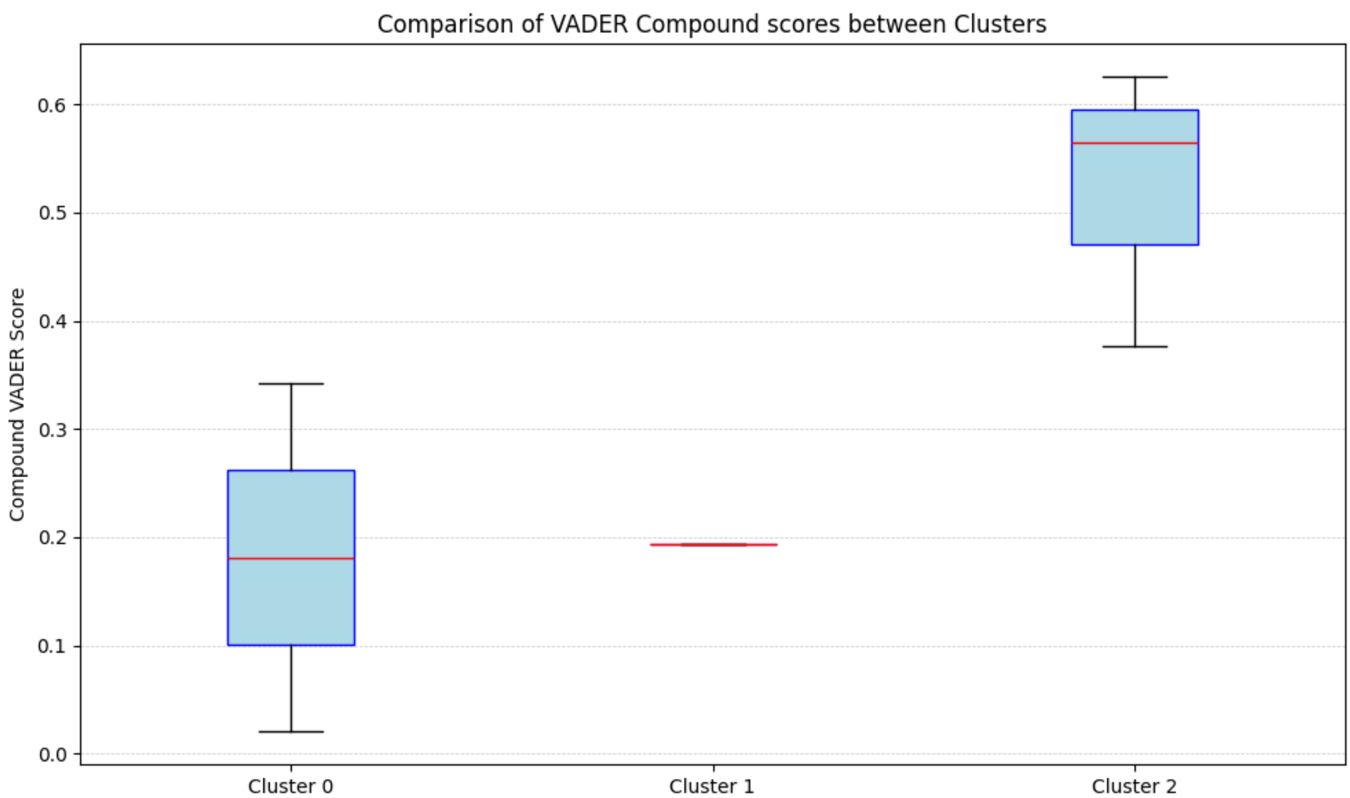

Figure 29: Comparison of VADER Compound scores between political Clusters

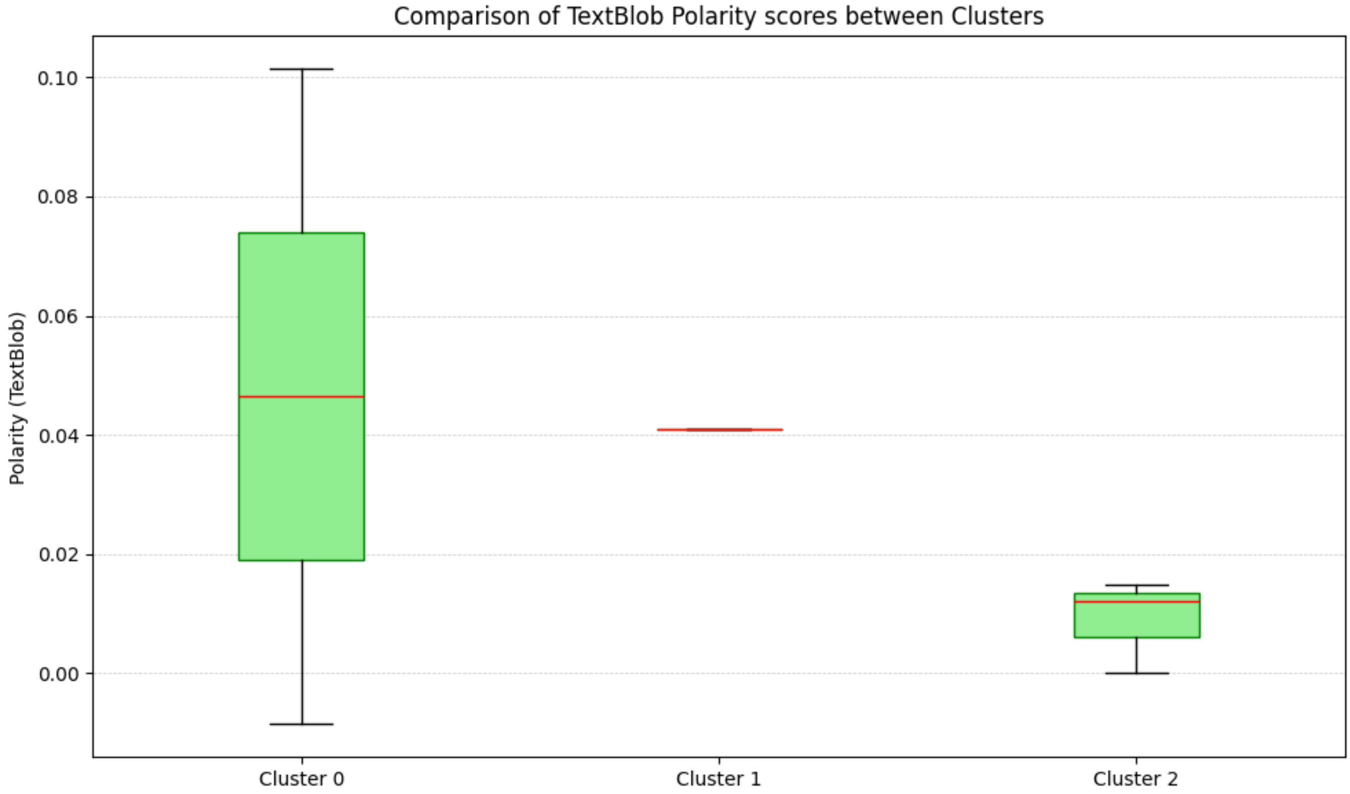

Figure 30: Comparison of TextBlob Polarity scores between political Clusters

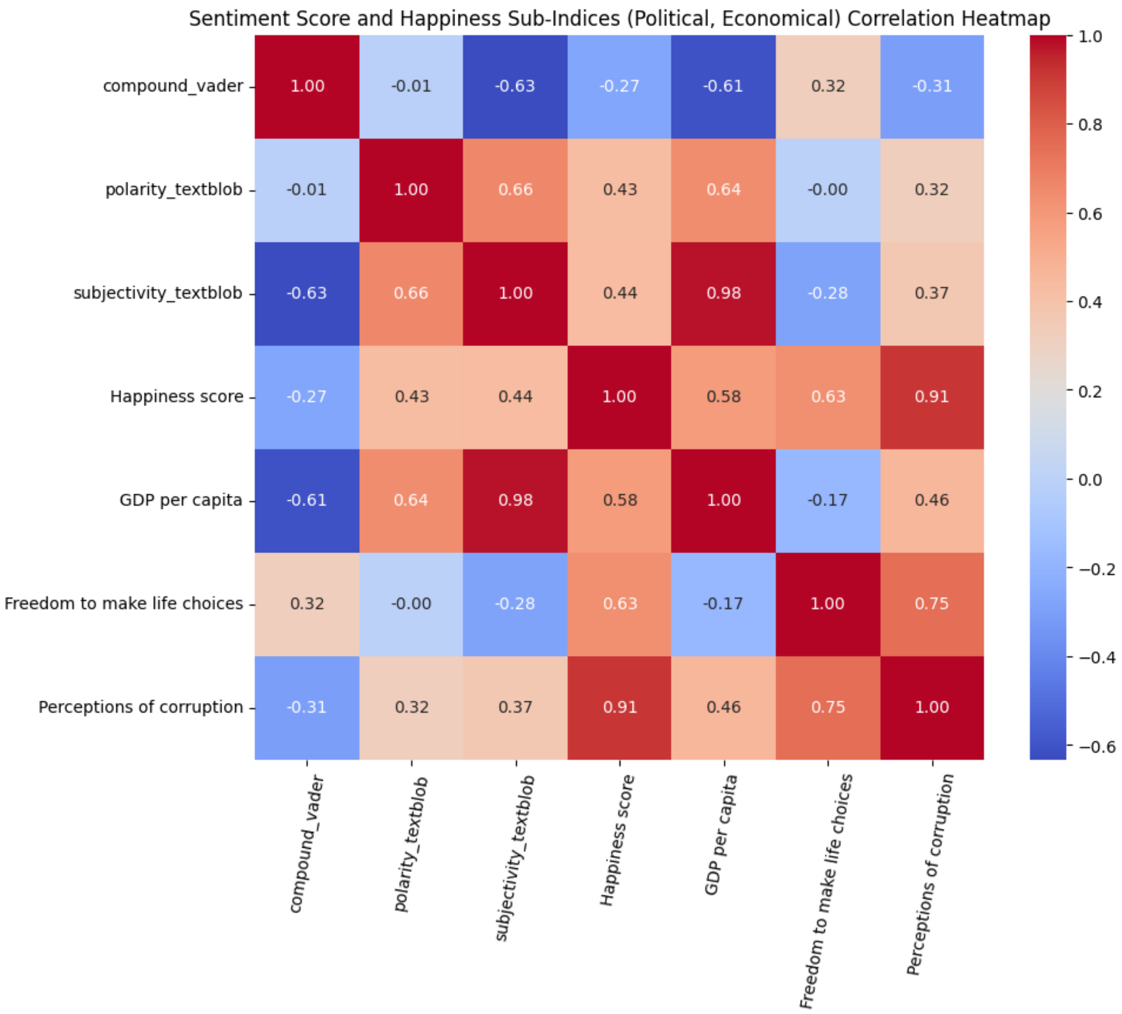

Figure 31: Sentiment Score and Happiness Sub-Indices (Political, Economical) Correlation Matrix

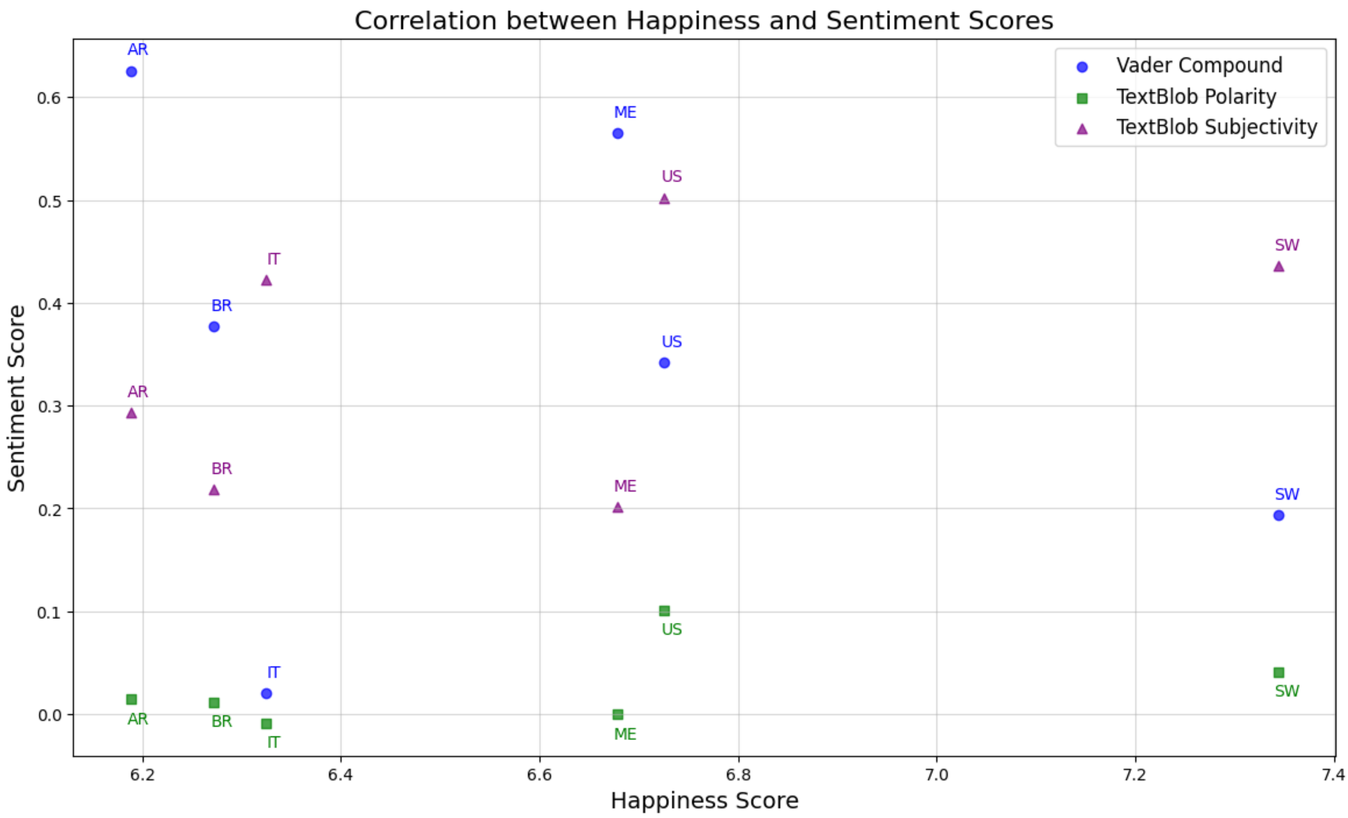

Figure 32: Comparison between Sentiment and Happiness Scores