### 

Twitter Sentiment Analysis: Spotify Tweet Analysis

Rochester Institute of Technology

Golisano College of Computing and Information Sciences

SaiTulasi Kamma / MS Data Science / [sk5656@rit.ed](mailto:sk5656@rit.edu)

**Executive Summary**

In the evolving landscape of digital music streaming, understanding public sentiment and engagement is crucial for companies to stay competitive and responsive to their users' needs. Spotify, a leader in this industry, operates not just as a service provider but as a cultural hub for music lovers worldwide. This report presents a comprehensive analysis of public perceptions and discussions related to Spotify, as derived from a dataset of tweets mentioning the brand.

The data, spanning various metrics and sentiment analysis, offers insights into how users perceive Spotify's service, engage with its features, and discuss its impact on their music consumption experience. By examining tweets and their associated sentiments, this study aims to uncover underlying trends and patterns that can inform strategic decisions and enhance user engagement strategies.

The following summary distills our findings from the tweet data, providing Spotify's stakeholders with actionable insights into the brand's social media presence and its implications for future business strategies. This analysis not only helps in pinpointing areas of strength but also identifies potential challenges within the public discourse, enabling Spotify to better align its offerings with the expectations and desires of its user base.

### **Data Preparation**

#### **Data Collection**

The initial dataset, titled "Spotify\_tweets\_2023.csv," comprises tweets that mention Spotify, capturing a wide range of discussions related to the music streaming service. This data was collected to understand public perceptions and to gauge sentiment trends associated with Spotify over time.

#### **Data Cleaning and Preprocessing**

The raw tweet data underwent several preprocessing steps to clean and standardize the text, making it suitable for detailed sentiment analysis and thematic exploration. The preprocessing steps included:

1. Noise Removal: The first stage of preprocessing involved removing various types of noise from the tweets to focus on meaningful text. This included:
   * Retweet Removal: Removing retweet tags to focus on original content.
   * Hashtag Removal: Stripping hashtags to prevent skewing text analysis with metadata.
   * URL Removal: Eliminating links to ensure the analysis reflects only textual content.
   * HTML Tag Removal: Cleaning up any embedded HTML tags that could distort data processing.
   * Mention Removal: Excluding mentions of other users to concentrate on the main message.
   * Carriage Return Removal: Removing newline and carriage return characters to clean up the entries.
   * Emoticon Removal: Stripping emoticons to standardize the text for linguistic analysis.

These steps were implemented using the tm package in R, applying transformations that strip unwanted characters and metadata from the tweets, thus refining the dataset for analytical purposes.

#### **Sentiment Analysis**

Post-cleaning, the dataset was subjected to sentiment analysis using two R packages: syuzhet and bing. These tools provided two sentiment scores per tweet, which were used to understand the emotional tone—positive, neutral, or negative—conveyed in the discussions about Spotify.

#### **Visualization**

To visualize the sentiment distribution and common themes, several plots were generated:

* Histograms: Created to show the frequency distribution of sentiment scores, helping identify the predominant sentiment in the data.**(Appendix Figure1, Figure 2)**
* Word Clouds: Used to highlight frequently occurring words, offering insights into the most talked-about aspects of Spotify.**(Appendix Figure3)**
* Keyword Analysis: Focused on the keyword "music" to explore specific discussions related to Spotify's core service area.

#### **Integration and Aggregation**

The cleaned and analyzed data was then aggregated to provide a comprehensive overview of sentiments over time, correlating these with specific events or promotional activities by Spotify if applicable.

### **Data Analysis & Results**

### **Tweet Volume and Engagement Patterns**

1. Volume Over Time: The graph titled "Time vs StatusId"**(Figure 4)** shows a significant spike in tweet volume in late April, indicating a major event or campaign that prompted increased Twitter activity. The volume sharply decreases thereafter, suggesting the end of a short-term event such as a product release or special promotion.
2. Text Width vs. Retweet(**Figure 16**): This scatter plot shows most tweets, regardless of text length, tend to have a relatively low retweet count, but there are a few outliers with high retweet counts. This indicates that while most tweets about Spotify do not gain significant traction, a few well-crafted or particularly resonant tweets achieve widespread sharing.

### **Sentiment Analysis**

1. Monthly Sentiment Trends: The line graphs indicating average sentiment scores (AvgSent1 and AvgSent2) **(Figure 7)** across months show variability:
   * AvgSent1: Shows a significant dip in April, suggesting a negative reception to whatever event occurred during this month. The sentiment improves and stabilizes in subsequent months.
   * AvgSent2: Mirrors the trends in AvgSent1 to a degree, but with less pronounced drops, indicating a slightly different measurement of sentiment that might be less sensitive to specific events.
2. Daily Sentiment in May**(Figure 6)**: The daily sentiment tracking for May reveals fluctuating but generally positive sentiment (AvgSent1) and more stable yet varied sentiment in AvgSent2. This could reflect ongoing discussions post-event with mixed reactions as users adjust to new updates or changes introduced in April.

### **Content Analysis**

1. Word Clouds**(Figure 12)** : These highlight the most frequently mentioned words in tweets. Terms like "music," "playlist," "premium," and "Spotify" dominate, indicating that most discussions are centered around Spotify’s core services. Other frequently appearing terms like "album" and "song" suggest that new music releases or specific content might be significant drivers of conversation.
2. Music vs. Retweet Volume**(Figure 9)**: A bar graph indicates that tweets associated with music are retweeted more often than those that are not. This suggests that Spotify's user base is highly engaged with content directly related to music, and such content is more likely to spread widely on the platform.

### **Correlations and Conclusions**

1. Sentiment vs. Favorites(**Figure 10**): The scatter plot demonstrates no strong correlation between the sentiment of a tweet and the number of favorites it receives, indicating that both positive and negative tweets can achieve high levels of engagement.
2. Engagement Insights(**Figure 11**): Despite the variability in sentiment and engagement, the data shows that strategic content related to music and Spotify's primary offerings tends to perform best in terms of reach and user interaction.

### **Strategic Recommendations**

* Leverage High-Engagement Topics: Focus on pushing content related to core themes like music, playlists, and premium services, which are proven to resonate with the audience.
* Monitor Sentiment Fluctuations: Given the volatility in sentiment observed, Spotify should continue to monitor public perception closely and adjust communications to address any negative trends.
* Capitalize on High-Retweet Opportunities: Identify characteristics of highly retweeted posts to enhance the effectiveness of future communications, potentially increasing reach and engagement.

This analysis provides a holistic view of how Spotify is discussed on Twitter, highlighting areas of strength and opportunities for improvement in public engagement and perception management.

### **Word Cloud**

The word cloud prominently features "Spotify" along with terms like "music," "song," "album," and "playlist." This highlights that discussions mainly revolve around Spotify's core offerings—music streaming and playlists. Other frequently mentioned words include "premium" and "account," suggesting conversations around subscription plans and user accounts. Additionally, brand-related tags like "Netflix" and "Amazon" indicate comparisons or integrations with other services, emphasizing Spotify's competitive and collaborative landscape in the digital entertainment sector.(**Refer Figure 12**)

### **Sentiment Over Time**

The "AvgSent1vsTime"(**Figure13, 15)** line graph illustrates the average sentiment score (sent1) of tweets across various months. Notably:

* Sentiment peaks in April, indicating particularly positive reactions or events in that month.
* A sharp decline is observed in May, followed by a recovery and stabilization in the following months. This pattern could reflect a response to specific updates, campaigns, or issues faced by Spotify in those months.
* The sentiment in October and November remains consistently high, suggesting positive public reception during these months, possibly influenced by new features, promotions, or seasonal content releases.

### **Sources of Tweets**

The bar chart depicting the "Count of Source" shows that the majority of Spotify-related tweets come from mobile devices, specifically "Twitter for iPhone" and "Twitter for Android," indicating that discussions about Spotify are predominantly mobile-driven. This aligns with Spotify's mobile-first user engagement strategy. The presence of "Twitter Web App" and "Sprinklr Care" as sources also suggests usage across different platforms, including professional monitoring and customer care responses via Sprinklr.(**Refer Figure 8, Figure 15**)

### **Monthly and Device Insights**

The monthly checkboxes and source filters in the visualization tools suggest an ability to drill down into more specific temporal or source-based analyses, which could be useful for understanding how sentiment and discussion topics vary across different times of the year or from different tweeting platforms.(**Figure 15**)

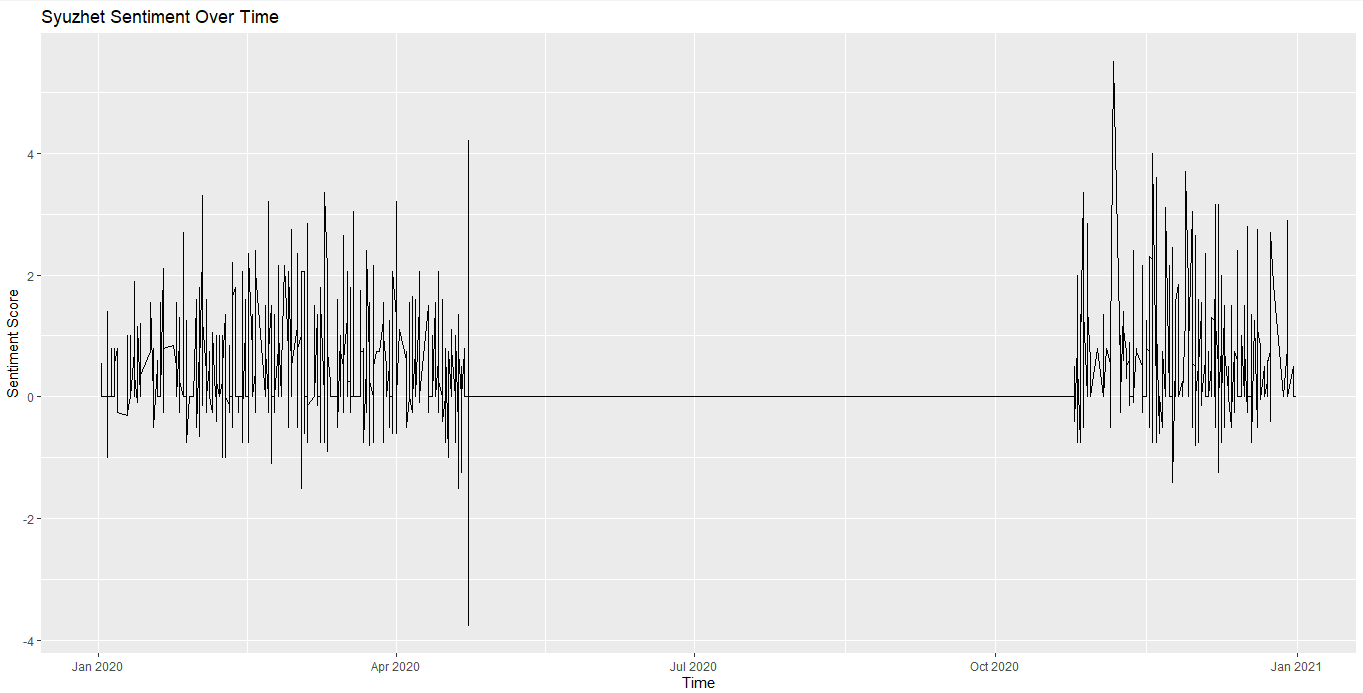
### **Interpretation**

This analysis provides valuable insights into how Spotify is discussed and perceived on social media:

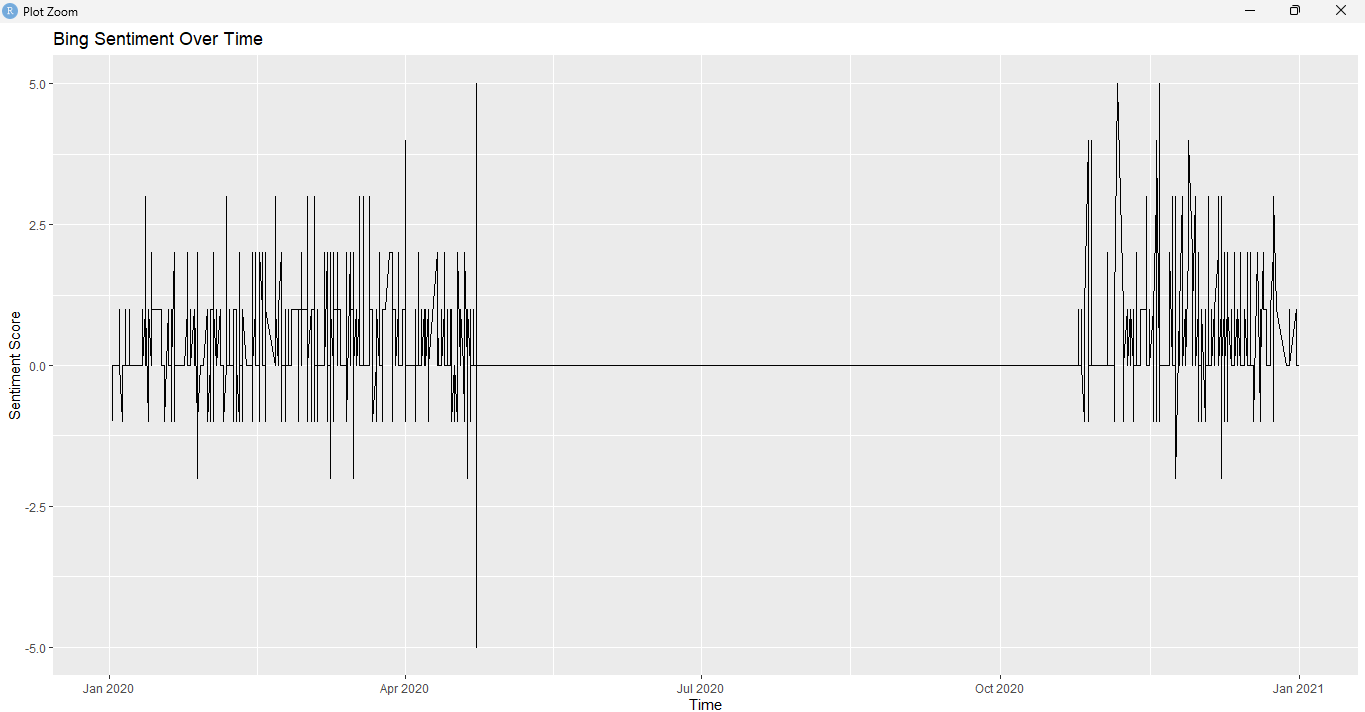
* The prevalent discussion topics revolve around Spotify’s service features, user experiences, and comparisons with other services.
* Sentiment analysis reveals fluctuations that could be tied to specific events, user experiences, or changes in service.
* The predominance of mobile sources for tweets underscores the importance of mobile platforms in user engagement strategies.

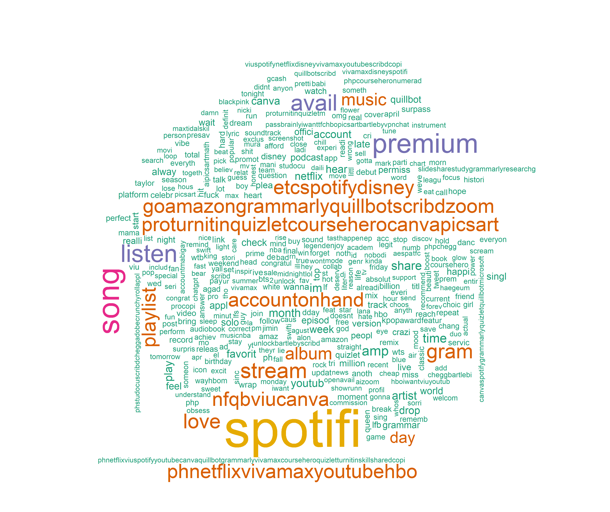
**APPENDIX**

**Figure 1**

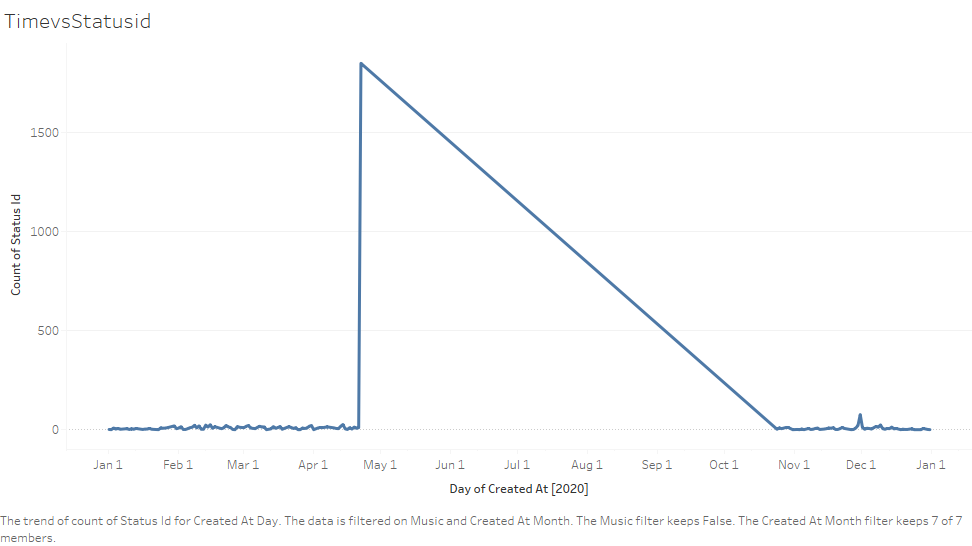
****

**Figure 2**

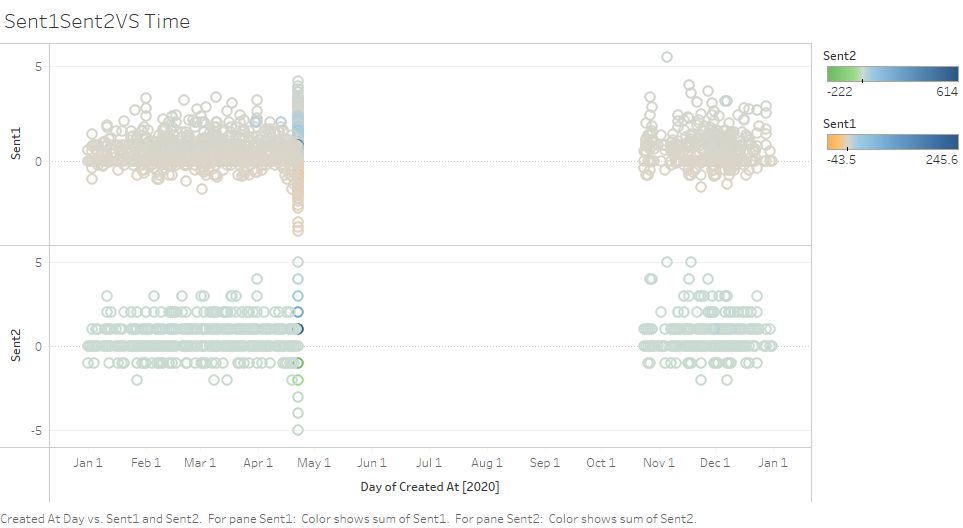
****

**Figure 3 Word Cloud Using R   
**

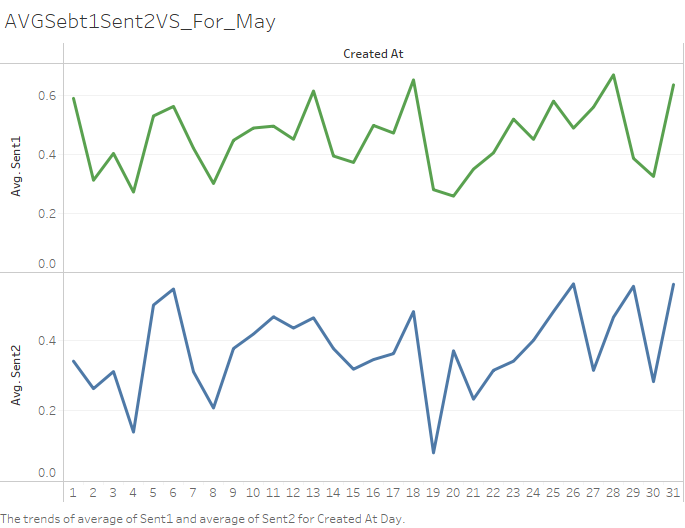
**Figure 4**

****

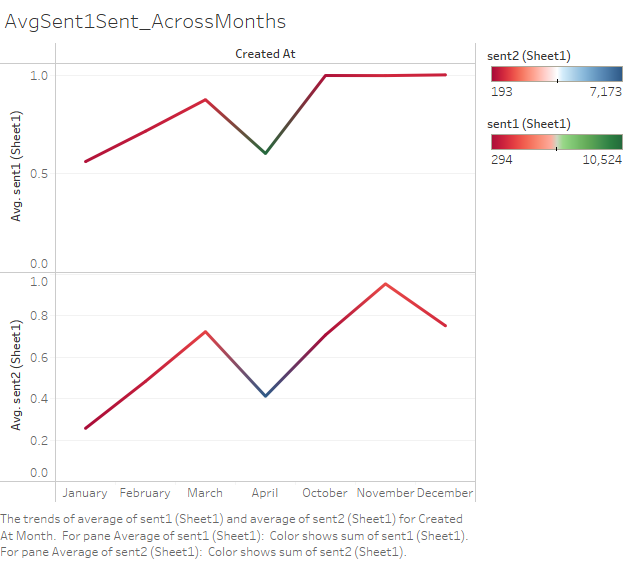
**Figure5**

****

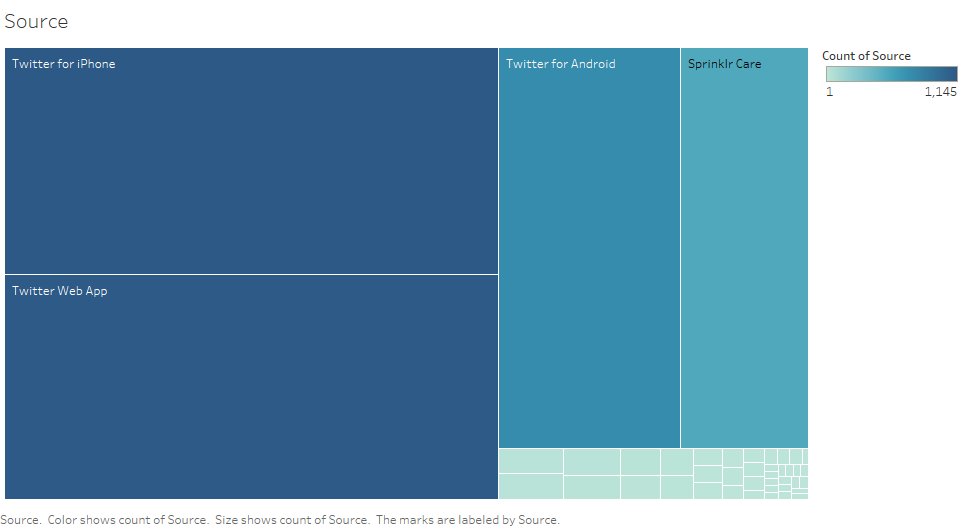
**Figure 6**

****

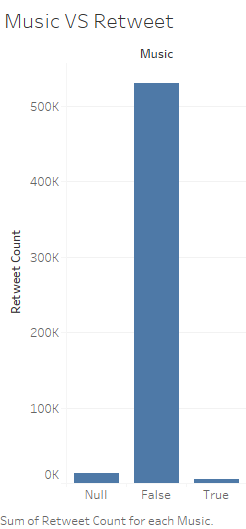
**Figure 7**

****

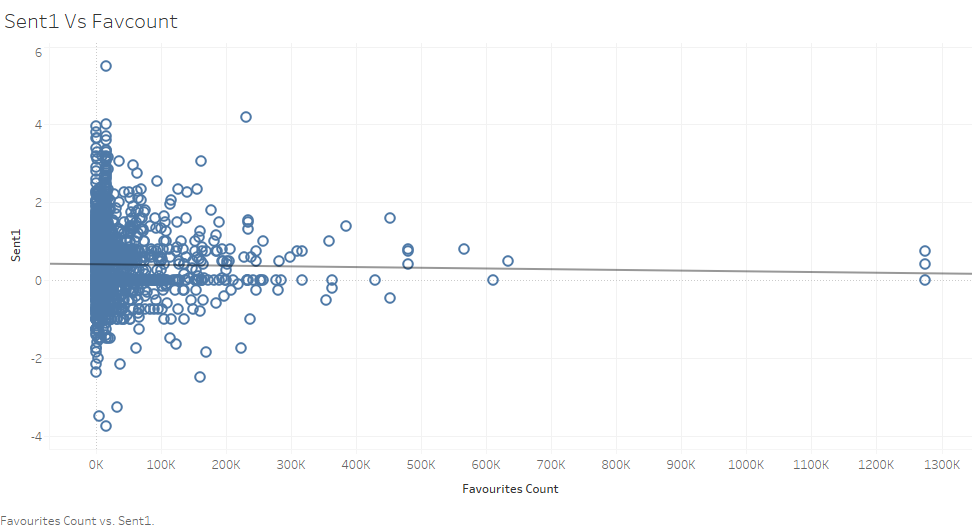
**Figure 8**

****

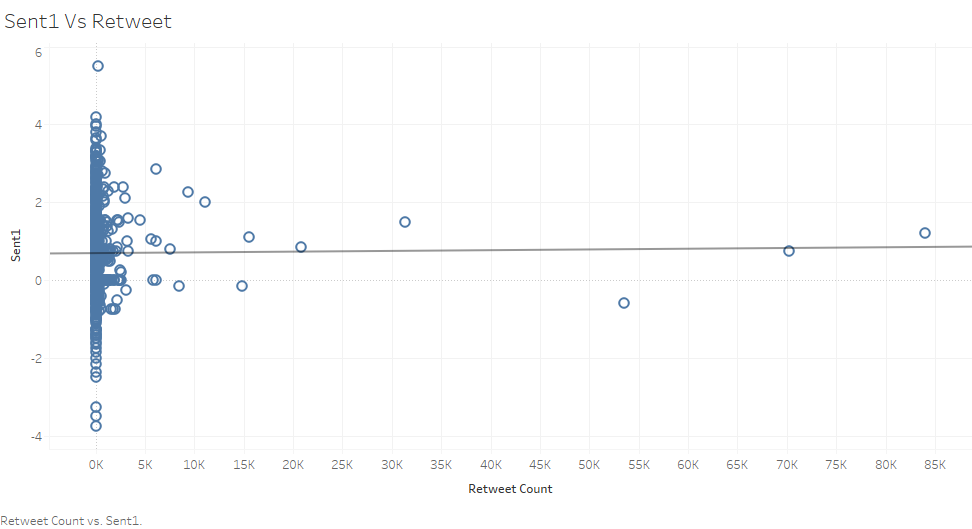
**Figure 9**

****

**Figure 10**

****

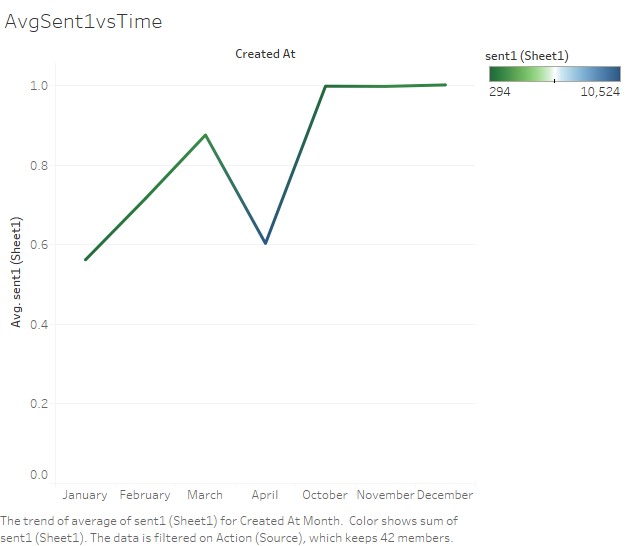
**Figure 11**

****

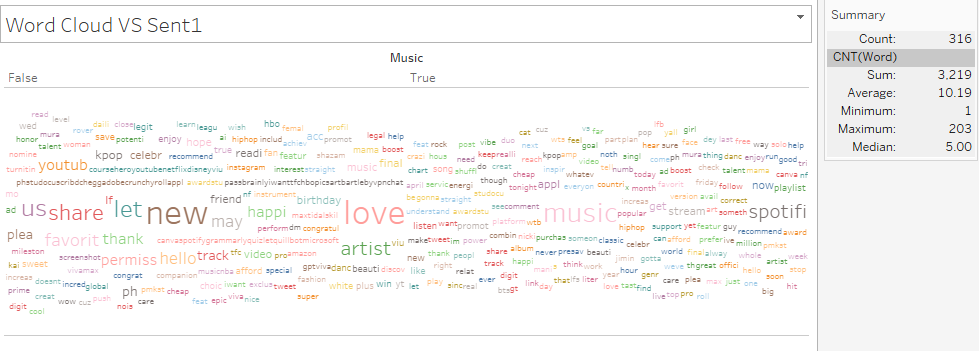
**Figure 12**

****

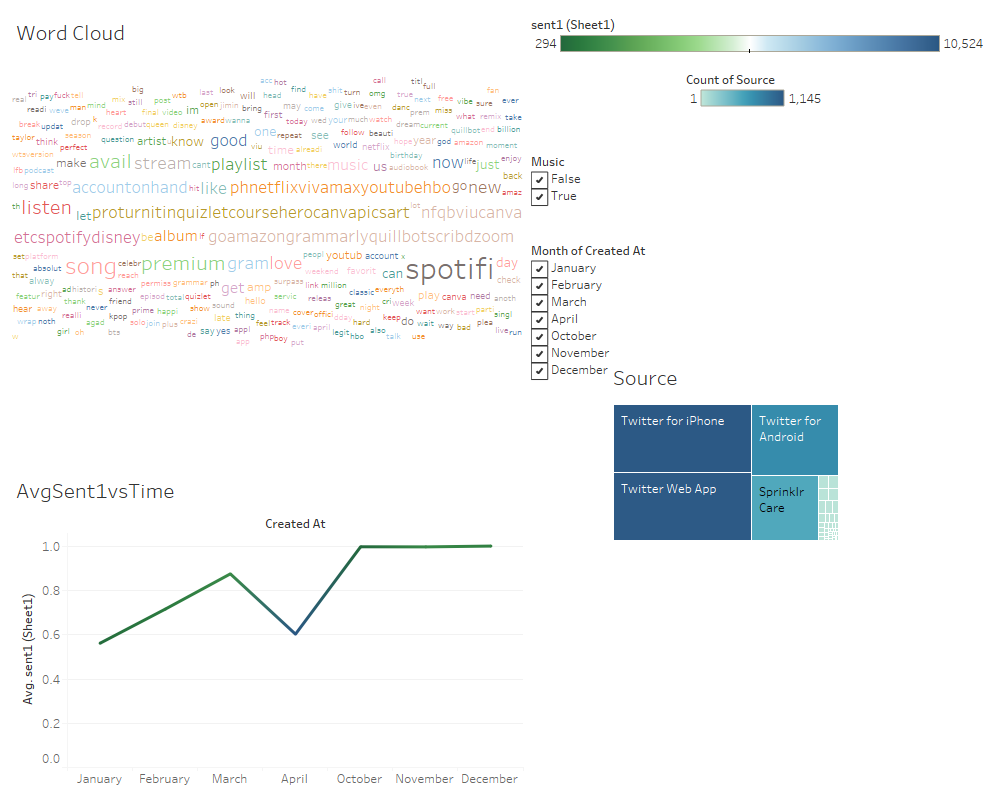
**Figure 13**

****

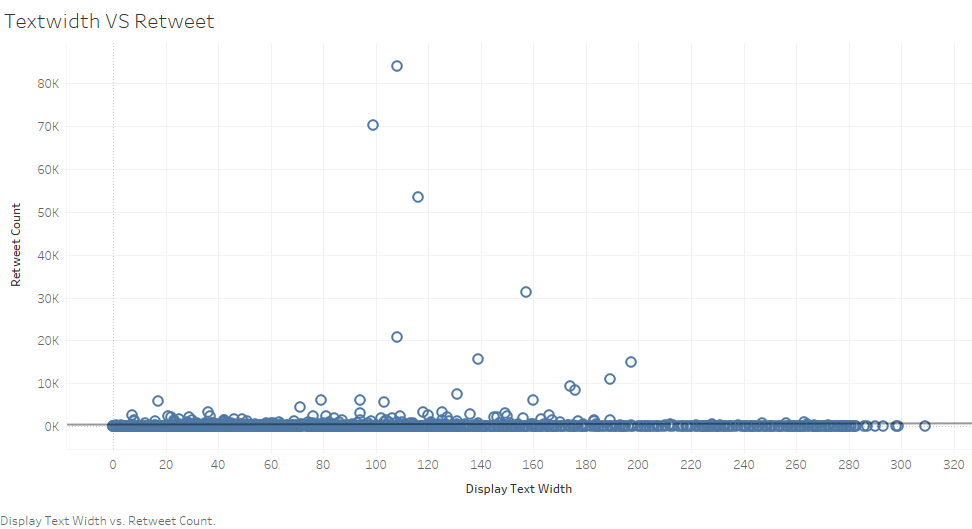
**Figure 14**

****

**Figure 15**

****

**Figure 16**

****