

Spotify Sentiment ✦ Analysis ✦



Dashboard Analysis

Piyush Korake (A032)

Partha Lalit (A034)

**Guided By:
Shweta Shirsat Mam**

RCDS 2024

Introduction

In the rapidly evolving landscape of digital entertainment, Spotify stands as a towering figure, offering a vast array of music and podcasts to users worldwide. Renowned for its expansive library and user-friendly interface, Spotify has become the premier destination for music enthusiasts and podcast aficionados alike. However, beneath its surface lies a wealth of user feedback and sentiments, waiting to be explored and understood.

This report embarks on a journey to dissect the sentiments expressed within Spotify reviews, utilizing a bespoke sentiment analysis dashboard crafted through a meticulous fusion of Python for sentiment analysis and Tableau for data visualization.

At the core of our exploration lies the sentiment analysis dashboard, a powerful tool designed to peel back the layers of user feedback and offer a comprehensive understanding of sentiments prevailing within the Spotify community. Through rigorous analysis and visual representation, we aim to distill key findings that provide invaluable insights into the overall sentiment towards Spotify and its diverse content offerings.

As we navigate through the dashboard, our quest is to uncover patterns, trends, and anomalies within the vast sea of Spotify reviews. From tracking sentiment trends over time to quantifying the distribution of sentiments across all reviews, our analysis traverses numerous dimensions, each offering a unique perspective into the nuanced landscape of user sentiments.

Furthermore, we delve into the correlation between user ratings and sentiments, seeking to unravel the intricate relationship between user perceptions and satisfaction levels. By deciphering this correlation, we aim to glean insights that can inform strategic decision-making and drive continuous improvement within the Spotify ecosystem.

Our journey culminates in the optimization of dashboard utilization, ensuring that stakeholders can effortlessly interpret analysis results and derive actionable insights. Equipped with intuitive visualizations and user-friendly interfaces, decision-makers can harness the power of user sentiment as a guiding force in shaping the future trajectory of Spotify.

In essence, this report serves as a testament to the transformative potential of sentiment analysis in unraveling the intricacies of user feedback within the digital realm. By leveraging advanced analytical tools and methodologies, we endeavor to empower stakeholders with the insights needed to drive meaningful enhancements and foster a more enriching user experience within the Spotify platform.

Objectives

1. **Sentiment Trends Monitoring:** Track fluctuations in sentiment over time to identify patterns and anomalies, providing insights into user satisfaction and potential issues with Spotify.
2. **Review Distribution Quantification:** Quantify the distribution of sentiments across all reviews to gauge overall user satisfaction, with a focus on identifying the proportion of positive, negative, and neutral reviews.
3. **Average Review Length Analysis:** Assess the average length of reviews for each sentiment category to understand user engagement levels and depth of experience with Spotify features.
4. **Rating-Sentiment Correlation Exploration:** Investigate the relationship between user ratings and sentiments to gain insights into how ratings influence user satisfaction and perception of Spotify.
5. **Dashboard Utilization Optimization:** Ensure efficient utilization of the sentiment analysis dashboard for quick interpretation of analysis results, providing stakeholders with clear and concise visualizations for informed decision-making based on user sentiment.
6. **Playlist and Chart Performance Comparison:** Analyze the song's performance across different streaming platforms (Spotify, Apple Music, Deezer) by comparing its presence in playlists and charts. Explore how the song's inclusion in various playlists and its chart rankings differ between platforms, and identify any correlations between playlist placements and chart positions.
7. **Music Characteristics Analysis:** Investigate the musical characteristics of the song, including tempo (bpm), key, mode, danceability, valence, energy, acousticness, instrumentalness, liveness, and speechiness. Determine how these attributes contribute to the song's popularity and identify any patterns or trends in the characteristics of successful songs.
8. **Streaming Metrics Correlation:** Explore correlations between streaming metrics such as total streams and playlist/chart presence across different platforms. Determine which factors have the strongest correlation with the song's streaming success and analyze how these metrics vary across different streaming services.
9. **Genre and Audience Analysis:** Utilize the song's characteristics and streaming metrics to identify its genre and target audience. Analyze how the song's genre and musical elements correlate with its streaming performance and audience engagement.

Data Description

	Time_submitted	Review	Rating	sentiment_scores	sentiment_category
0	09-07-2022 15:00	great music service audio high quality app eas...	5	{'neg': 0.0, 'neu': 0.411, 'pos': 0.589, 'comp...	positive
1	09-07-2022 14:21	please ignore previous negative rating app sup...	5	{'neg': 0.276, 'neu': 0.267, 'pos': 0.458, 'co...	positive
2	09-07-2022 13:27	popup get best spotify experience android 12 a...	4	{'neg': 0.148, 'neu': 0.495, 'pos': 0.357, 'co...	positive
3	09-07-2022 13:26	really buggy terrible use recently	1	{'neg': 0.458, 'neu': 0.542, 'pos': 0.0, 'comp...	negative
4	09-07-2022 13:20	dear spotify get songs didnt put playlist shuf...	1	{'neg': 0.0, 'neu': 0.583, 'pos': 0.417, 'comp...	positive
...
61589	01-01-2022 03:01	even though communicated lyrics feature availa...	1	{'neg': 0.0, 'neu': 0.828, 'pos': 0.172, 'comp...	positive
61590	01-01-2022 02:13	use sooo good back downloaded free version cou...	1	{'neg': 0.181, 'neu': 0.471, 'pos': 0.347, 'co...	positive
61591	01-01-2022 01:02	app would good taking device start comp theres...	2	{'neg': 0.283, 'neu': 0.566, 'pos': 0.151, 'co...	negative
61592	01-01-2022 00:49	app good hard navigate wont let play song clic...	2	{'neg': 0.18, 'neu': 0.659, 'pos': 0.161, 'com...	positive
61593	01-01-2022 00:19	good sometimes doesnt load music plays like 8 ...	4	{'neg': 0.129, 'neu': 0.503, 'pos': 0.368, 'co...	positive

61594 rows × 5 columns

To start the analysis, we collected the Spotify Reviews dataset from Spotify Sentiment Analysis. The dataset contains customer reviews and associated information for the popular music streaming service. We loaded the dataset into Python and performed initial data exploration to understand its structure and features. The dataset includes attributes such as the time the review was submitted, the text within the review, the rating (1-5), the total thumbs up (how many people thought the review was helpful), and the review reply.

- Time Submitted: The date and time when each review was posted.
- Review: The actual text of the review left by the user.
- Rating: The numerical rating given by the user, likely on a scale of 1 to 5.
- Sentiment Scores: A breakdown of the sentiment analysis, showing the proportions of negative, neutral, and positive sentiments, along with a compound score.
- Sentiment Category: The overall sentiment of the review categorized as positive or negative.

track_name	artist(s)_name	artist_count	released_year	released_month	released_day	in_spotify_playlists	in_spotify_charts	streams_in_spotify	in_apple_playlists	in_apple_charts	in_deezer_playlists	in_deezer_charts	in_shazam_charts	bpm	key	mode	danceability_%	valence_%	energy_%	acousticness_%	instrumentalness_%	liveness_%	speechiness_%
Seven (feat. J. Balvin, Jung Kook)	Seventeen	2	2023	7	14	553	147	1E+08	43	263	45	10	826	125	B	Major	60	89	83	31	0	8	4
LALALA	Myke Towers	1	2023	3	23	1474	48	1E+08	48	125	58	14	382	32	C#	Major	71	61	74	7	0	10	4
vampire	Olivia Rodrigo	1	2023	6	30	1397	113	1E+08	94	207	91	14	949	138	F	Major	51	32	53	17	0	31	6
Cruel Summer	Taylor Swift	1	2019	8	23	7858	100	8E+08	116	207	125	12	548	170	A	Major	55	58	72	11	0	11	15
WHERE SHE GOES	Bad Bunny	1	2023	5	18	3133	50	3E+08	84	133	87	15	425	144	A	Minor	65	23	80	14	63	11	6
Sprinter	Dave Navarro	2	2023	6	1	2186	91	2E+08	67	213	89	17	946	141	C#	Major	52	66	58	19	0	8	24
Elle Balla So	Elisabon Arr	2	2023	3	16	3090	50	7E+08	94	222	43	13	418	148	F	Minor	67	83	76	48	0	8	3
Columbia	Quevedo	1	2023	7	7	714	43	6E+07	25	89	30	13	194	100	F	Major	67	26	71	37	0	11	4
Lukumen	Gunn	1	2023	5	15	1096	83	1E+08	60	210	48	11	953	130	C#	Minor	85	22	62	12	0	28	5
La Bala - Re Pasa Pluma	La Bala	2	2023	3	17	2953	44	6E+08	45	10	66	13	335	170	D	Minor	91	86	48	21	0	8	33
un x1000	Bad Bunny, J	2	2023	4	17	2876	40	5E+08	41	205	54	12	251	83	F#	Minor	57	56	72	23	0	27	5
Super Shy	NewJeans	1	2023	7	7	422	58	6E+07	37	202	21	5	188	150	F	Minor	78	52	82	18	0	15	7
Flowers	Miley Cyrus	1	2023	1	12	12211	119	1E+09	300	25	745	58	1021	118		Major	71	65	68	6	0	3	7
Daylight	David Kushner	1	2023	4	14	3538	38	4E+08	80	156	182	24	1251	130	D	Minor	51	32	43	83	0	9	3
As It Was	Harry Styles	1	2022	3	31	23575	130	3E+09	403	198	863	46		174	F#	Minor	52	66	73	34	0	31	6
Kill Bill	Drake	1	2022	12	8	8109	77	1E+09	183	162	151	12	187	89	G#	Major	64	43	73	5	17	16	4
Dupid - Twi	Fifty Fifty	1	2023	2	24	2942	77	5E+08	91	212	78	6	0	120	B	Minor	78	76	58	43	0	34	3
What Was I Thinking	Billie Eilish	1	2023	7	13	873	104	3E+07	80	227	95	24	1173	78		Major	44	14	9	96	0	10	3
Classy 101	Feid, Young	2	2023	3	31	2610	40	3E+08	43	100	54	14	187	100	B	Major	86	67	66	14	0	12	16
Like Crazy	Jinny	1	2023	3	24	596	68	4E+08	8	104	23	2	25	123	G	Major	63	36	73	0	0	36	4
LADY GAGA	Gabbri Baller	3	2023	6	22	332	28	8E+07	11	163	10	4	0	140	F	Minor	65	87	74	22	0	42	4
I Can See Yo	Taylor Swift	1	2023	7	7	516	38	5E+07	73	118	42	1	150	123	F#	Major	69	82	76	6	0	6	3
Wanna Be	Arctic Monke	1	2013	1	1	12853	110	1E+09	24	98	582	2	73	135		Minor	48	44	42	12	2	11	3
Peso Pluma	Bisnarp, Pe	2	2023	5	31	1313	40	2E+08	17	162	32	11	128	123	F	Minor	85	81	67	26	0	12	5
Regular (feat. The Weeknd)	Drake	2	2023	6	2	1945	87	1E+08	74	182	87	14	1093	99	C#	Major	88	83	68	7	0	36	20

- track_name: Name of the song
- artist(s)_name: Name of the artist(s) of the song
- artist_count: Number of artists contributing to the song
- released_year: Year when the song was released
- released_month: Month when the song was released
- released_day: Day of the month when the song was released
- in_spotify_playlists: Number of Spotify playlists the song is included in
- in_spotify_charts: Presence and rank of the song on Spotify charts
- streams: Total number of streams on Spotify
- in_apple_playlists: Number of Apple Music playlists the song is included in
- in_apple_charts: Presence and rank of the song on Apple Music charts
- in_deezer_playlists: Number of Deezer playlists the song is included in
- in_deezer_charts: Presence and rank of the song on Deezer charts
- in_shazam_charts: Presence and rank of the song on Shazam charts
- bpm: Beats per minute, a measure of song tempo
- key: Key of the song
- mode: Mode of the song (major or minor)
- danceability_%: Percentage indicating how suitable the song is for dancing
- valence_%: Positivity of the song's musical content
- energy_%: Perceived energy level of the song
- acousticness_%: Amount of acoustic sound in the song
- instrumentalness_%: Amount of instrumental content in the song
- liveness_%: Presence of live performance elements
- speechiness_%: Amount of spoken words in the song

Dashboard Creation

In the beginning we collected the data from online resources. As the sentiment analysis consists of text data we used Python to analyze the data with the help of BERT model and generated the sentiment scores for the data set.

We used Tableau to import the data set and analyze it. As the data was categorical, to convert it in desired format we created necessary calculated fields such as review length and count of sentiments.

Our first visualization is the line graph of month vs count of sentiments which gave us sentiment over time line graph for all three sentiments i.e. positive, neutral and negative respectively. To make it more interactive and informative we added filters to choose between any of the three sentiments and another filter to choose days between the daily range over 01 Jan to 09 July 2022.

The second visualization is the sentiment distribution of count of review and sentiment score. Here we can observe the difference of count of each sentiment through bar chart.

Next we visualized the average sentiment review length by it's respective category. Here we could analyze the length of text written by reviewer to express their sentiment.

As the sentiments are also associated with its ratings we also visualized a subdivided bar plot between rating and count of reviews. Here we also used the sentiment category to understand the ratio of all the three sentiments with respect to it's rating. It helps us in understanding the customer reviews more deeply.

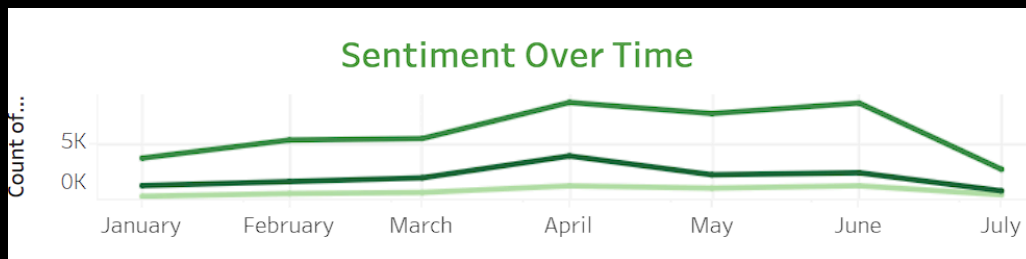
Lastly we plotted a box plot to visualize the central tendency, and variability of a dataset in a concise visual format.

In the end we created a dashboard combining all the plots and to make them more informative we applied filters used in the first viz to all other visualizations.

This completes our dashboard creation procedure.

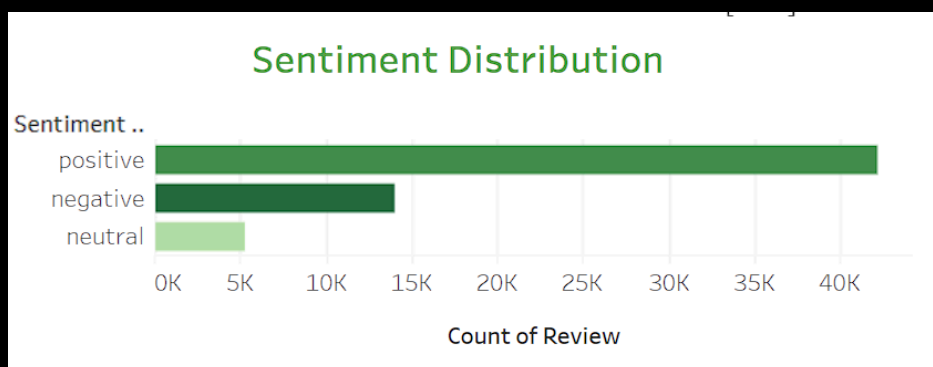
Results and Interpretation

1. Sentiment Over Time



We can observe that positive sentiment peaks in April and June with 8683 and 8631 count of reviews respectively and starts lowering after June. While the negative sentiment also peaks in April with 3896 counts it is less than positive and the neutral sentiment count is close to 1000 through out the year.

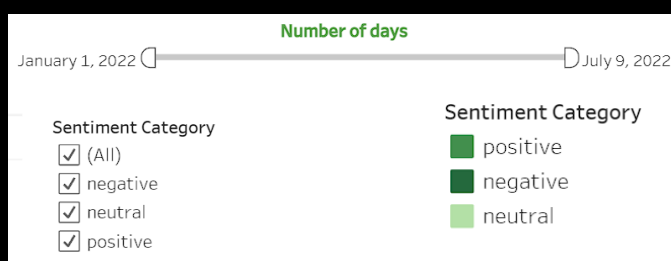
2. Sentiment Distribution



The positive sentiment is maximum followed by negative and neutral. This concludes that more customers are satisfied with the service.

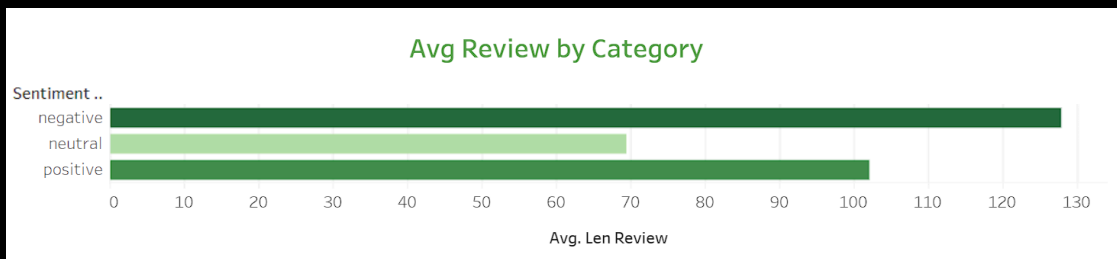
Positive review counts: 42182, Negative review counts: 14066, Neutral review counts: 5336

3. Filters



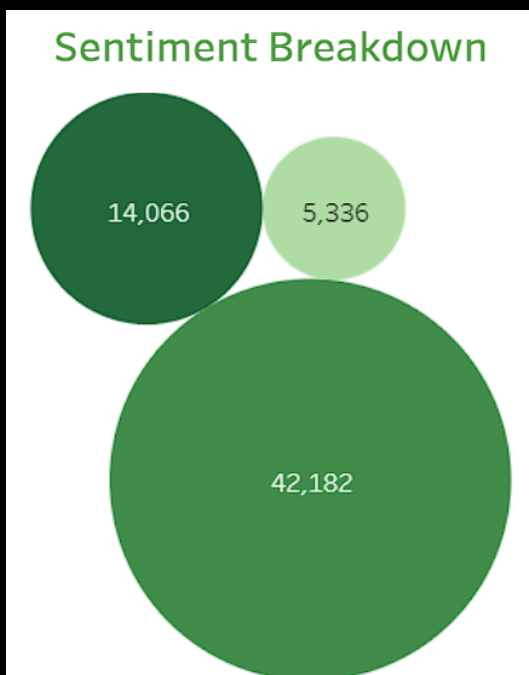
The slider helps in adjusting the dates in selected range to change the dashboard respectively. The sentiment category checkbox displays the selected sentiments and the color index represents the color for each sentiment in entire dashboard respectively.

4. Average Review Length by Category



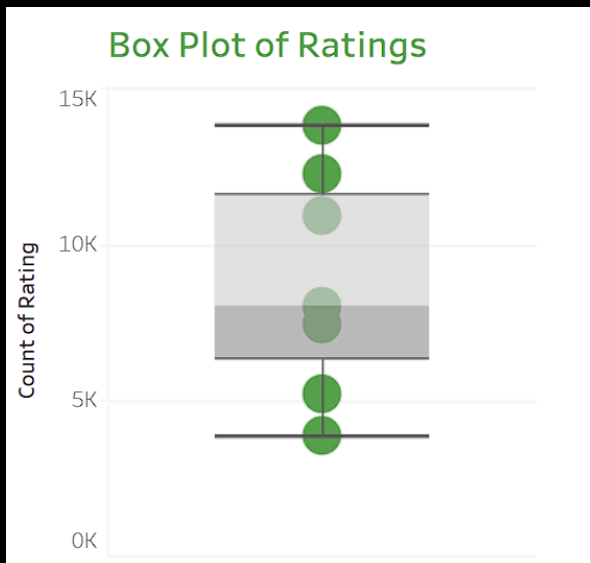
The average review length is maximum for negative followed by positive and neutral respectively. This suggests that the negative reviews are written in detail. Positive avg review length: 102.14 , Negative avg review length: 127.88 , Neutral avg review length: 69.59

5. Sentiment Breakdown



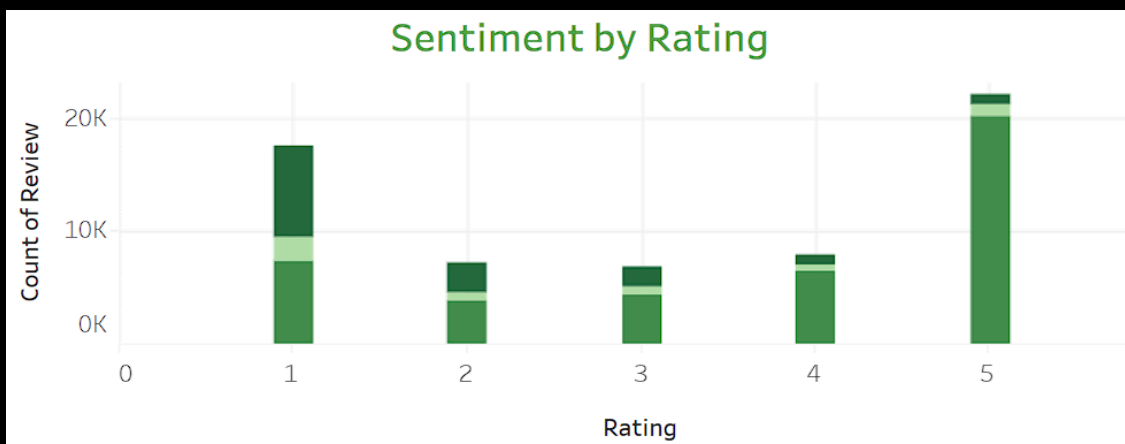
The bubble chart suggests that the positive sentiment is maximum followed by negative and neutral sentiment respectively.

6. Box Plot of Ratings



The presence of outliers on the both sides of boxplot suggests that the data may have extreme values in both the end of distribution. The median being close to lower whisker suggests that the data is positively skewed. March month is median with the count of 8023 reviews and April and July are outliers at extreme ends with review counts 13803 and 3900 respectively.

7. Sentiment by Rating

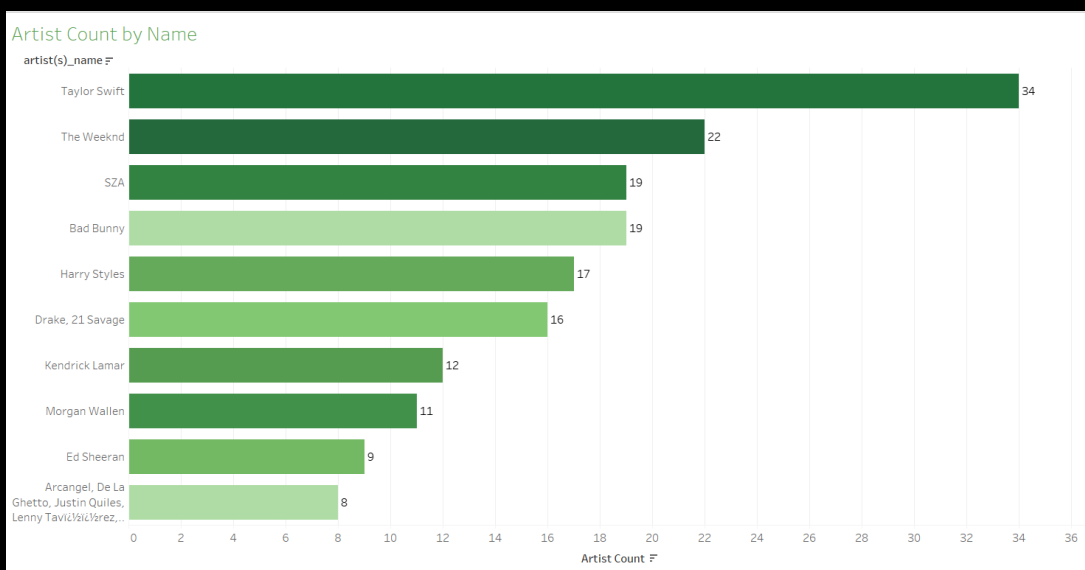


This subdivided plot suggests that the positive review sentiment has the maximum count in the 5 rating.

The the negative sentiment has got maximum rating as 1 but also contains positive sentiment in it. Refer to table given below

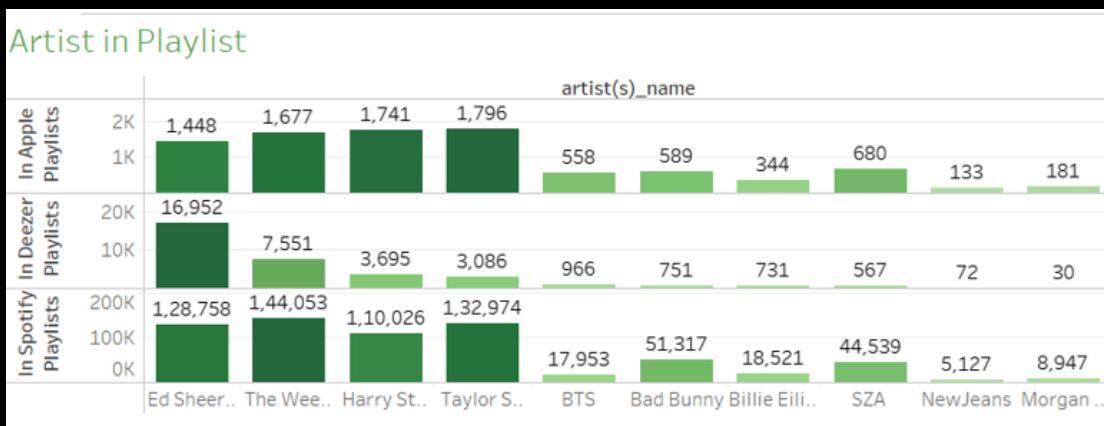
Rating	Positive Sentiment Count	Neutral Sentiment Count	Negative Sentiment Count
1	7301	2200	8150
2	3823	780	2515
3	4434	711	1741
4	6422	570	850
5	20202	1075	810

8. Artist count by name



The bar graph suggests that Taylor Swift has the maximum number of songs released with the count of 34 followed by The Weekend and SZA

9. Artists in Playlist



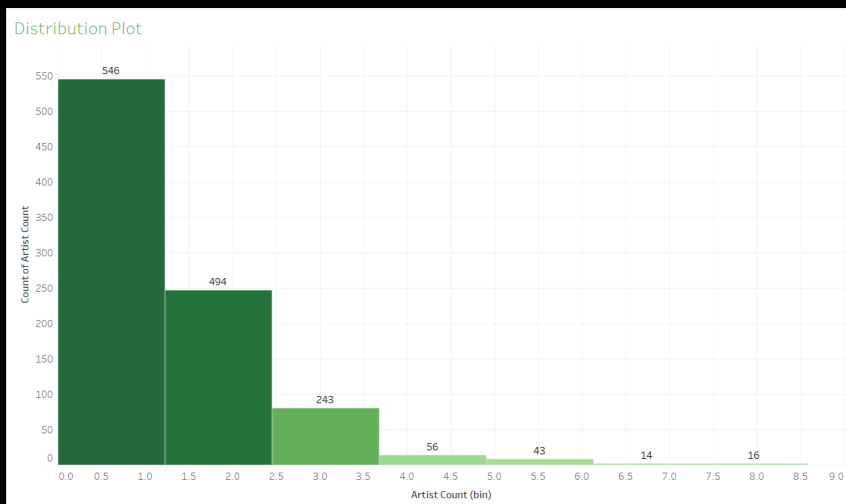
We can clearly observe that,

Apple users prefer listening to Taylor Swift and Harry Styles with the count of 1796 and 1741 respectively.

Deezer users listen to Ed Sheeran with the count of 16952

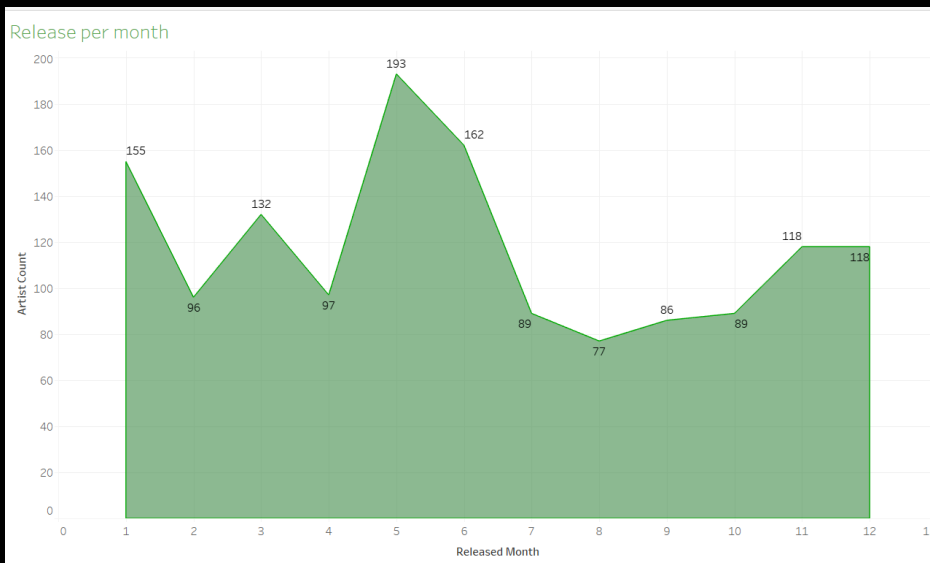
Spotify users prefer The Weeknd has max listeners followed by Taylor Swift and Ed Sheeran with counts 144053, 12394 and 128758 respectively.

10. Distribution of Artist Count



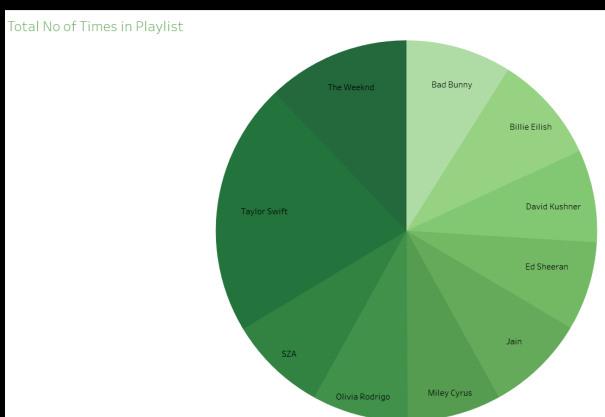
We can observe that maximum songs composed by at most 2 artists are 1040 and only 16 songs were composed by 8 artists.

11. Releases per Month



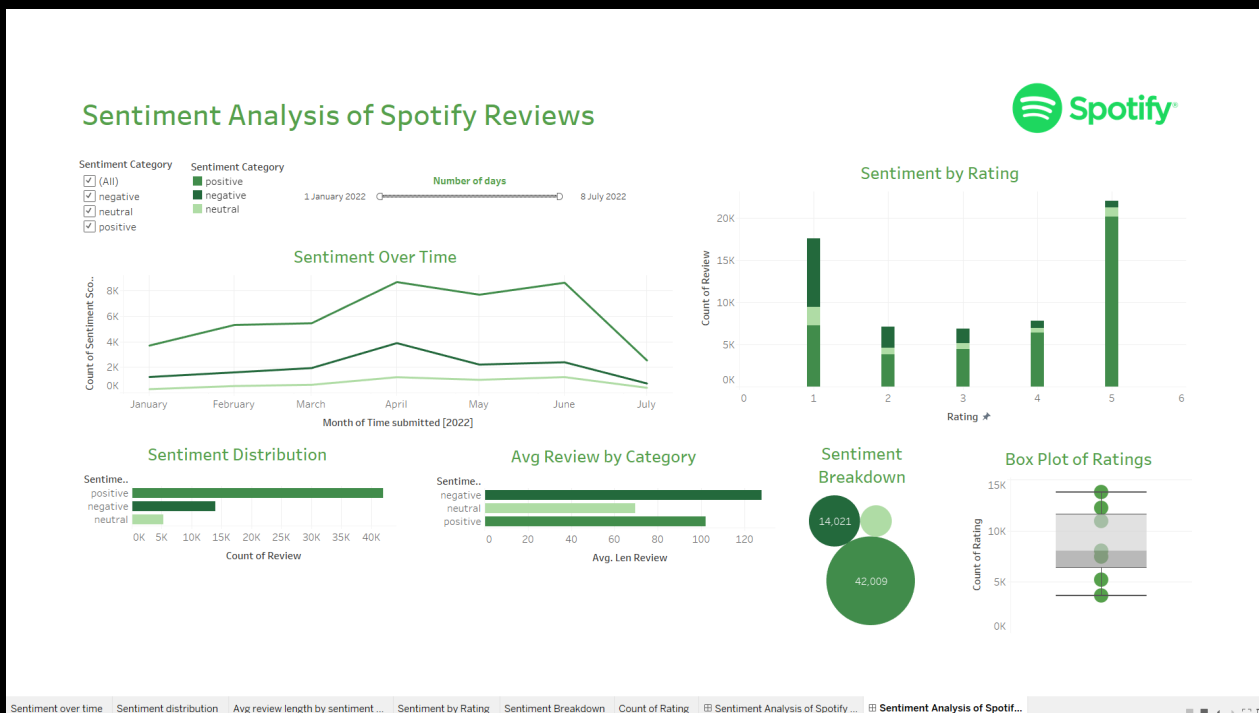
We can observe that the maximum releases of the song happen in the mid-year 193,162 counts respectively followed by the beginning and end of the year.

12. Pie chart of Artist's Frequency

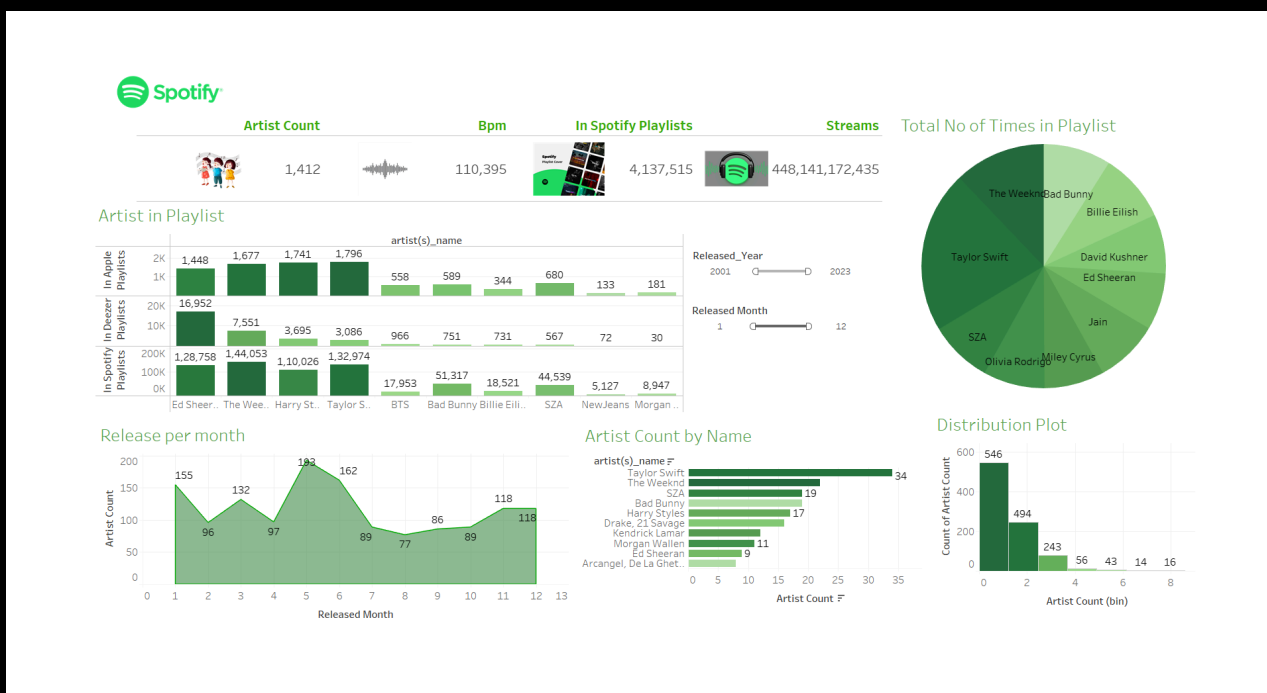


The pie chart shows the frequency share of each artist signifying Taylor Swift to hold maximum share and least hold is of Ed Sheeran

Dashboard



Dashboard 1



Dashboard 2

The dashboard can be accessed at:

[Tableau Dashboard1](#)

[Tableau Dashboard2](#)

The data referred data set can be accessed at:

[Data1 Kaggle](#)

[Data2 Kaggle](#)

Conclusion

The analysis conducted illuminates Taylor Swift's indomitable presence in the music sphere, epitomized by her staggering volume of releases and steadfast popularity across diverse streaming platforms. The sheer magnitude of Swift's catalog, alongside the formidable positions of artists like The Weeknd and SZA, underscores the evolving landscape of contemporary music, where individuality and creativity reign supreme. Platform-specific preferences unearth intriguing insights into the varied tastes of listeners, painting a nuanced portrait of audience behavior across platforms like Apple Music, Deezer, and Spotify. Furthermore, the prevalence of collaborative efforts in music production hints at a collaborative spirit within the industry, fostering innovation and cross-pollination of creative energies. Meanwhile, the discernible peaks in song releases throughout the year suggest a strategic undercurrent in scheduling and marketing practices, reflecting a dynamic ecosystem driven by calculated moves and artistic expression.

Future Scope

Moving forward, deeper exploration into platform-specific preferences and collaborative dynamics promises to unlock a wealth of untapped insights into audience behavior and industry trends. By harnessing advanced sentiment analysis techniques and augmenting datasets with additional attributes, such as demographic information and geographical nuances, a more comprehensive understanding of user sentiments and preferences can be attained. Furthermore, ongoing refinement and enhancement of sentiment analysis dashboards will ensure that stakeholders remain equipped with actionable insights to drive continuous improvement in user experience and platform engagement. Embracing a proactive approach to monitoring and adapting to evolving user sentiments will be instrumental in fostering sustained growth and innovation within the ever-evolving landscape of the music industry.