

# Big Data Analytics

Vinyl Revisited: Analyzing Market Dynamics and Consumer Sentiments

Kent Business School| Diviyaj Ayengia

## Table of Contents

### Introduction to Data and Research Question

#### 1.1 Research Justification

#### 1.2 Research Methodologies

#### 1.3 Research Question Development

### Data Processing & Exploration

#### 2.1 Variables Available

#### 2.2 Selection and Transformation

#### 2.3 Data Exploration

### Results and Data Visualization

#### 3.1 Popularity Trends in Music Genres

#### 3.2 Consumer Sentiments Analysis

#### 3.3 Correlations between Album Features and Consumer Behaviors

### Discussion

#### 4.1 Impact of Genre on Consumer Preferences

#### 4.2 Strategic Implications for the Music Industry

### Conclusion

#### Summary of Findings

#### Recommendations for Stakeholders

### References

## 1. Introduction to Data and Research Question

The music industry's evolution has been marked by the striking resurgence of vinyl records alongside the continued presence of CDs, embodying a blend of nostalgic allure and audiophile preference in an era dominated by digital formats. This revival, often termed a "renaissance," is not merely a cultural re-emergence but reflects significant shifts in technology and consumer psychology (Guo, 2023). Vinyl's comeback has been robust, attracting both older aficionados and new generations of listeners who appreciate its unique sonic qualities. Komara (2015) highlights this trend among younger consumers who favour the tangible and auditory superiority of vinyl over digital alternatives. Additionally, López Medel (2014) examines the symbolic role of album covers, suggesting that vinyl's physical attributes contribute to its cultural and collectible value, enhancing its appeal in the digital age.

## 1.1 Research Justification

The resurgence of vinyl records, despite the dominance of digital music platforms, underscores a significant shift in consumer behaviour and market dynamics within the music industry. The vinyl revival is not only about nostalgia but also represents a choice for audio quality and physical interaction with music media. Guo (2023) discusses this phenomenon in the context of vinyl's historical and contemporary appeal, noting a "renaissance" in the digital era that reflects changes in technology and consumer psychology.

Additionally, the physical attributes of music products, particularly vinyl, play a crucial role in consumer purchase decisions. López Medel (2014) explores the symbolic function of album covers in the vinyl revival, emphasizing their transition from mere packaging to elements of popular culture that enhance the consumer's experience. Komara (2015) further supports this by highlighting the resurgence of vinyl among younger consumers who value its tangible and superior audio qualities over digital formats. These studies underscore the relevance of examining how different music formats are perceived in terms of consumer sentiment and purchasing behaviour, which is crucial for understanding the current market trends and future directions of the music industry.

This study aims to extend the utility of sentiment analysis into this underexplored trend. Sentiment analysis is a powerful tool widely utilized for prediction across diverse sectors, demonstrating significant impact in domains such as stock markets, elections, and consumer behaviour. Examples include its use for forecasting stock trends (Nguyen et al., 2015; Mittal & Goel, 2012), analysing electoral outcomes (Ramteke et al., 2016; Budiharto & Meiliana, 2018), and predicting e-commerce product pricing (Tseng et al., 2017). It has also been effective in anticipating future buying behaviours (Yadav et al., 2019) and evaluating service ratings in sectors like hospitality (Barbosa et al., 2015) and the entertainment industry (Jain, 2013). However, its application within specific niches like music genres in the CDs and Vinyl sector has been less investigated, particularly in how consumer sentiments derived from reviews predict genre popularity and preferences.

Employing sentiment analysis techniques to explore consumer reviews within the "CDs and Vinyl" dataset. This involves identifying general sentiments—positive, negative, or neutral—and correlating them with music genre popularity, as highlighted by the literature indicating the value of sentiment analysis in various predictive contexts (Guo, 2023; Komara, 2015).

- Examining the distribution of sentiments across different music genres to derive insights into consumer preferences, supported by the analysis of consumer sentiments in niche markets

like vinyl records, where physical and sensory attributes significantly influence consumer behaviour (López Medel, 2014).

- Investigating the features of genres that garner more attention or higher ratings, potentially indicating broader popularity or consumer preferences. This includes a detailed examination of the linguistic elements and common phrases in consumer reviews, using bigrams and other text analysis tools to gauge the emotional tone and thematic consistency (Nguyen et al., 2015; Komara, 2015).
- Utilizing machine learning models and natural language processing tools to conduct a nuanced exploration of how sentiments expressed in consumer reviews correlate with trends and preferences in music genres. This approach builds on the established methodologies and aims to provide a more focused application of sentiment analysis in predicting music genre popularity within the "CDs and Vinyl" category (Guo, 2023).

The study's objective is not only to enrich our understanding of consumer behaviour in the music market but also to enhance predictive modelling techniques for music genre popularity based on detailed sentiment analysis. This approach ensures a comprehensive integration of sentiment analysis into a sector where it has not been extensively applied, leveraging it to uncover insights that are critical for market analysis and strategic planning in the music industry.

## 1.2 Research Methodologies

### Data Collection

The primary data for this research was sourced from Amazon's "CDs and Vinyl" category, which includes extensive consumer reviews, ratings, and product metadata. This dataset provides a rich basis for analysing consumer preferences and behaviours.

### Data Processing

The initial data processing involved:

- *Data Merging*: Combining review data with metadata based on product ASINs to ensure comprehensive data points for each record.
- *Data Cleaning*: Standardizing and cleaning the text data, converting data types appropriately, and handling missing values to prepare the dataset for analysis.

### Exploratory Data Analysis (EDA)

- *Genre Categorization*: Music genres were categorized and analysed to determine their popularity and trends over time.
- *Sentiment Analysis*: Consumer reviews were classified as positive, negative, or neutral based on their ratings. Text analysis was further conducted to extract sentiments from review texts, using natural language processing techniques to understand deeper consumer sentiments.

### Network Analysis

Employed to explore and visualize the relationships within the data, particularly focusing on the interaction between positive and negative sentiments within the reviews.

This research utilizes a comprehensive methodological framework to address the specified research questions, employing both quantitative and qualitative analyses to explore consumer behaviour in the CDs and Vinyl market. The methodologies are supported by literature that underscores the cultural and economic significance of understanding consumer preferences in a rapidly evolving digital landscape.

### 1.3. Research Question Development

The study aims to delve deep into the dynamics of consumer preferences within the CDs and Vinyl categories using a dataset comprising Amazon reviews, which includes ratings, text reviews, and product metadata. The research will address the following questions:

- Which genres within the CDs and Vinyl category are most popular, and how has their popularity evolved over recent years?
- How do consumer sentiments towards these music genres vary, and what factors contribute to the positive or negative sentiments?
- How do album features and metadata correlate with consumer purchasing behaviours and review ratings within specific music genres?

## 2. Data Processing & Exploration

### 2.1 Variables Available

The dataset from Amazon's "CDs and Vinyl" category is comprehensive, containing various types of data including review texts, ratings, metadata like genre, price, and release dates. For our analysis, key variables used were:

- *Genre*: Essential for understanding the variety and distribution of music genres within the dataset.
- *Review Ratings*: Provides a quantitative metric of consumer satisfaction and popularity.
- *Review Texts*: Qualitative insights are gleaned from these texts, offering deeper understanding of consumer sentiments.
- *Category Information*: Used to further classify music into more specific sub-genres or themes, aiding in detailed trend analysis.
- *ASIN*:

### 2.2 Selection and Transformation

Given the complexity and volume of the data, the following transformations were necessary:

- *Cleaning Review Texts*: Removed irrelevant characters and standardized text formatting for consistent input in sentiment analysis.
- *Genre Categorization*: Genres were grouped into broader categories to simplify analysis and focus on significant trends rather than granular distinctions.
- *Rating Normalization*: Standardized ratings across different sub-datasets to ensure consistency in measuring consumer satisfaction.
- *Category Split*: Additional classification was performed on the genre data to isolate specific trends within sub-genres, making the analysis more relevant to niche market trends.

### 2.3 Data Exploration

Exploratory data analysis was conducted through a series of visualizations to identify patterns and relationships within the data:

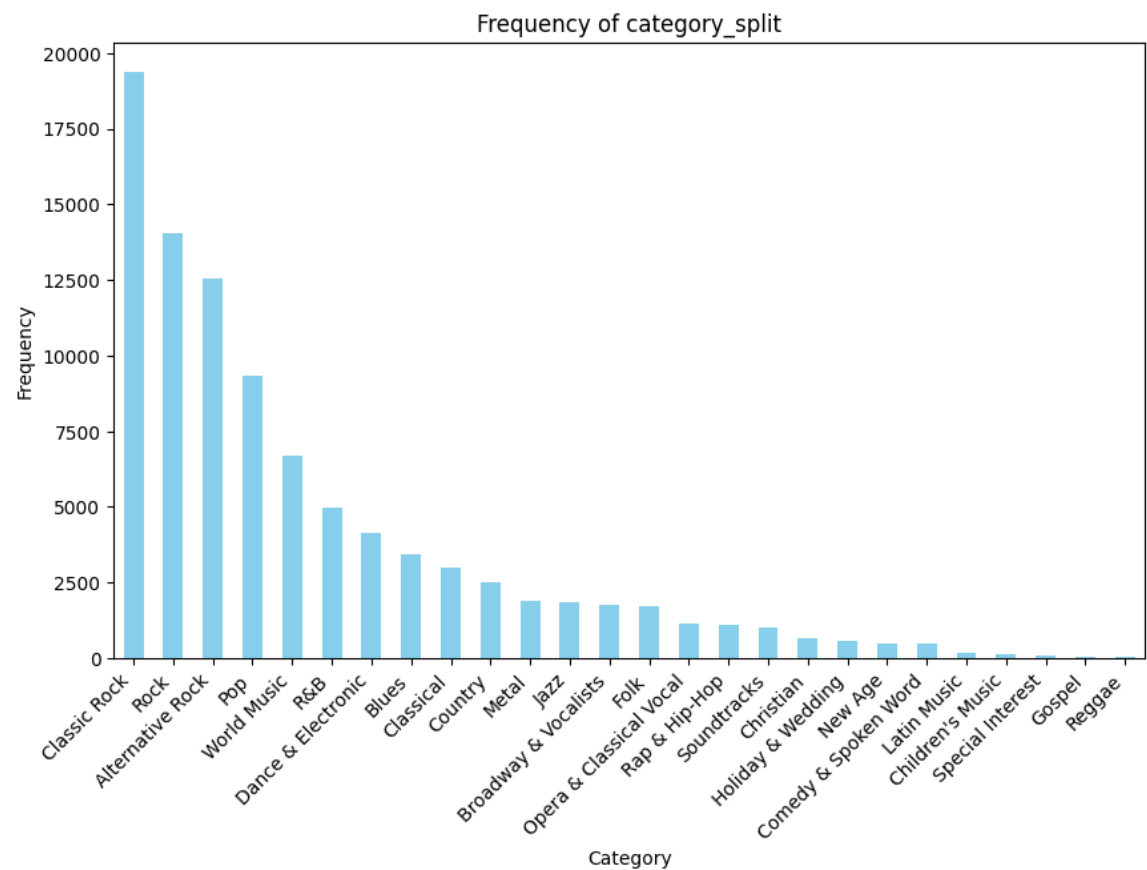
*Genre Popularity:* Visualizations indicated a clear consumer preference for genres such as rock and pop, evident from the volume of reviews and average ratings.

*Sentiment Analysis:* Text analysis of review contents helped extract sentiments associated with different genres, providing insights into consumer satisfaction levels.

*Category Analysis:* Detailed analysis of sub-genres revealed nuanced trends within broader categories, highlighting specific areas of consumer interest and potential market opportunities.

### 3. Results and Data Visualization

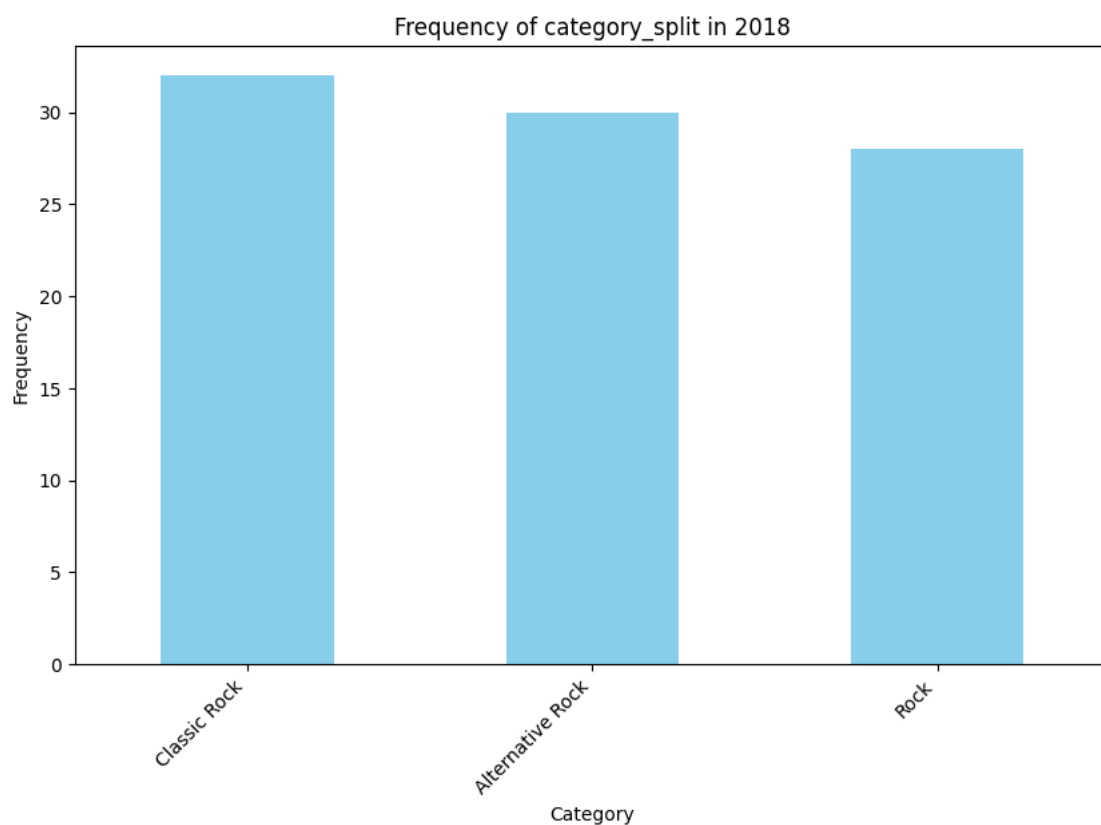
3.1 Which genres within the CDs and Vinyl category are most popular, and how has their popularity evolved over recent years?

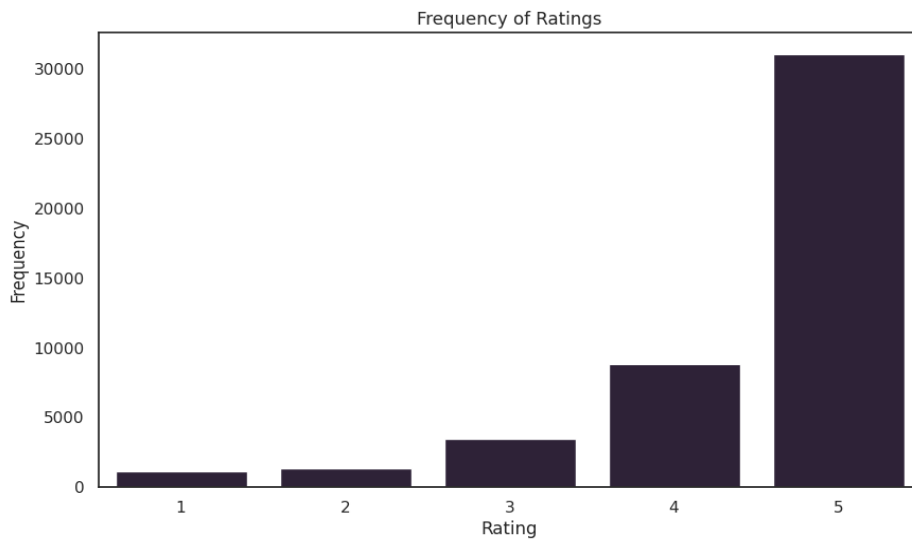


The bar chart provided illustrates the frequency of distribution of music genres most mentioned within the "CDs and Vinyl" category, clearly highlighting the dominance of Classic Rock, Rock, and Alternative Rock genres. These genres notably outpace others in popularity, with Classic Rock leading significantly, suggesting a robust affinity among consumers for these music styles. Following these, World Music and Pop also show considerable popularity, though they do not reach the heights of the rock categories. The lower frequency of genres such as Metal, Jazz, and Classical further down the scale indicates niche markets within the broader music category. It is evident from the chart that rock music, particularly

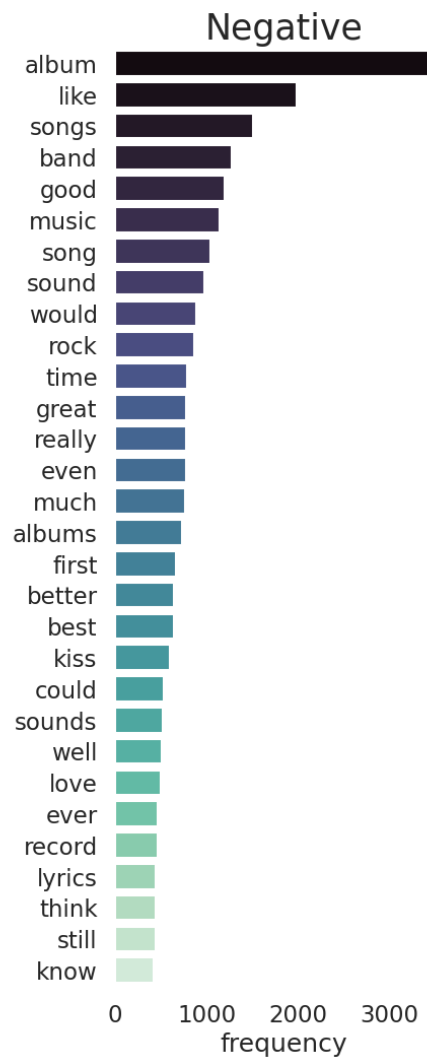
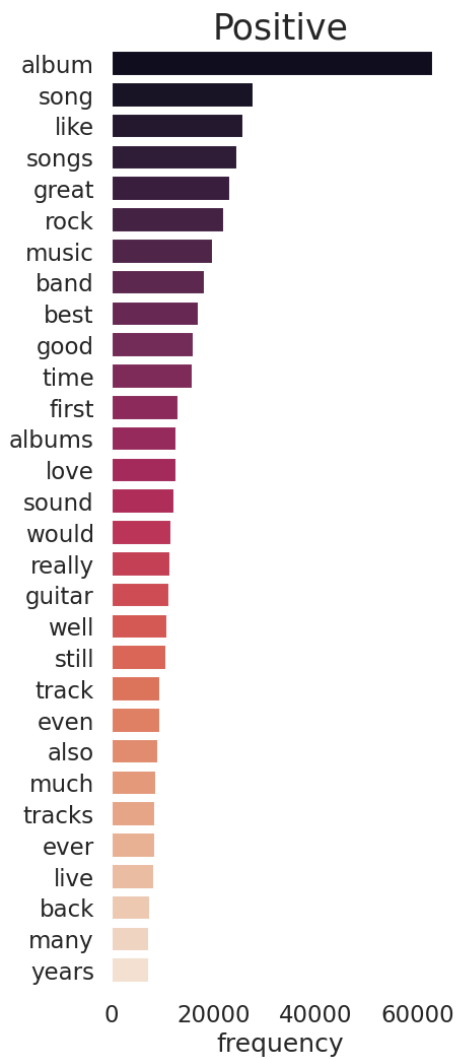
Classic Rock, maintains a commanding presence in the music market, aligning with consumer preferences for enduring classics and influential musical styles. This trend likely reflects both a nostalgic attachment to the music of past decades and a continuous interest in the genre's evolution, underscored by the substantial numbers of consumers actively purchasing these genres.

The bar chart titled "Frequency of category\_split in 2018" illustrates the distribution of reviews across rock genre for the year 2018, specifically focusing on Classic Rock, Alternative Rock, and Rock. Each of these genres exhibits a comparable level of frequency in terms of reviews, ranging around 25 to 30 reviews per genre for the year. this graph provides a snapshot for 2018, suggesting that rock music, in its various forms, continues to hold substantial popularity. It implies that in 2018, there was no significant variance in preference between these styles of rock music, indicating a stable consumer interest in rock genres during this period. This can be a useful insight when looking at trends over time to determine if the popularity of these genres is increasing, decreasing, or remaining stable.





3.2 How do consumer sentiments towards these music genres vary, and what factors contribute to the positive or negative sentiments?





### *Common Words in Positive Reviews:*

The terms "album," "song," "great," "best," and "love" dominate in positive reviews. These words suggest that aspects like overall album composition, individual song quality, and emotional connection are significant drivers of positive sentiments among consumers.

The presence of words like "first," "still," "back," and "years" indicates that nostalgia or long-standing appreciation for certain albums or bands may contribute to positive reviews.

### *Common Words in Negative Reviews:*

Negative reviews also frequently mention "album" and "song," which indicates that dissatisfaction often revolves around specific aspects of the music itself.

Words like "better," "best," "could," and "think" suggest that expectations are not met, with consumers comparing their experience to potential or previous standards.

The word "kiss" stands out and might need further context to understand its specific relevance; it could be related to a particular album or song that received mixed reviews.

The graph provides empirical evidence on the vocabulary predominantly associated with positive and negative sentiments in reviews. This can help identify which features or aspects

of music are most appreciated or critiqued, such as quality of music, emotional impact, and expectations.

#### *Genre-Specific Analysis:*

By examining the context in which these words are used across different genres, you can discern if certain genres are more prone to positive or negative reviews based on these attributes. For example, genres like Classic Rock might see positive sentiments associated with "love" and "best" due to nostalgic value, whereas more contemporary genres might attract criticism that hinges on "better" or "could," indicating a gap between expectations and reality.

#### *Marketing and Product Development:*

Understanding which words frequently appear in reviews can guide marketing strategies and product development. Highlighting positively associated words in promotional materials can resonate more with potential buyers. Similarly, addressing commonly critiqued aspects mentioned in negative reviews can improve product offerings and customer satisfaction.

#### *Enhanced Customer Insights:*

This analysis helps in tailoring communication and customer service by prioritizing aspects that consumers care about most. For instance, if "sound quality" frequently appears in negative contexts, focusing on improving and communicating enhancements in this area could mitigate negative perceptions.

#### *Conclusion:*

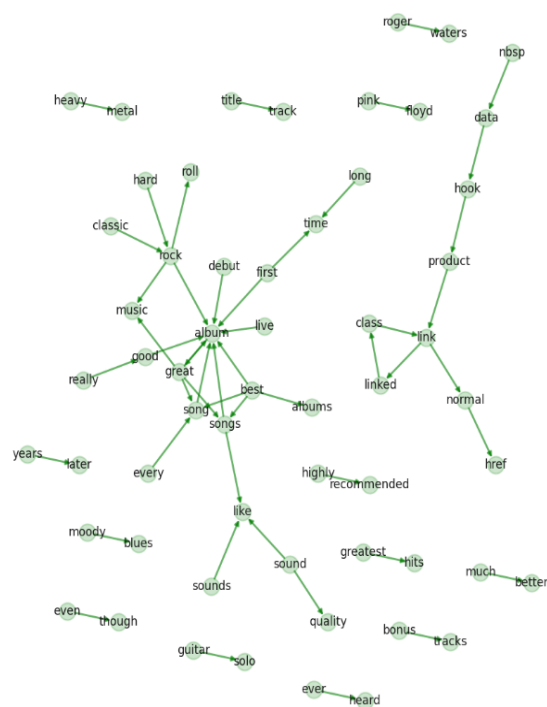
This word frequency graph is an invaluable tool for understanding consumer behavior and sentiments towards CDs and Vinyl. It helps address your research question about how sentiments vary across music genres and what factors influence these sentiments, enabling targeted actions to enhance consumer satisfaction and strategically position products within the market.

## **Bigram**

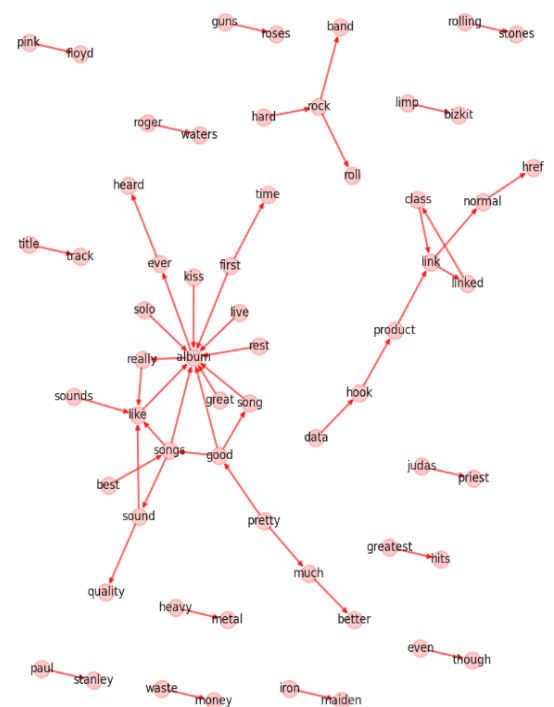
This visualization of positive and negative review networks is pivotal to our research as it elucidates the specific attributes that drive consumer satisfaction and dissatisfaction within the CDs and Vinyl category. By mapping out the associations between key terms in both positive and negative contexts,

we gain invaluable insights into what consumers cherish and criticize in their music purchases. For example, the positive connections around terms like "great," "album," and "classic rock" highlight the attributes that resonate with and fulfil consumer expectations, suggesting areas for potential focus in marketing and product development. Conversely, the negative associations reveal critical pain points and areas for improvement, such as disappointment with certain albums or artists. Understanding these dynamics allows for a more informed approach to addressing consumer needs and enhancing their overall experience, which is essential for maintaining relevance and competitiveness in the evolving music industry. This analysis not only aids in tailoring products to meet consumer tastes but also informs strategic decisions that can help mitigate negative feedback and reinforce positive brand perceptions.

## Positive Reviews



## Negative Reviews

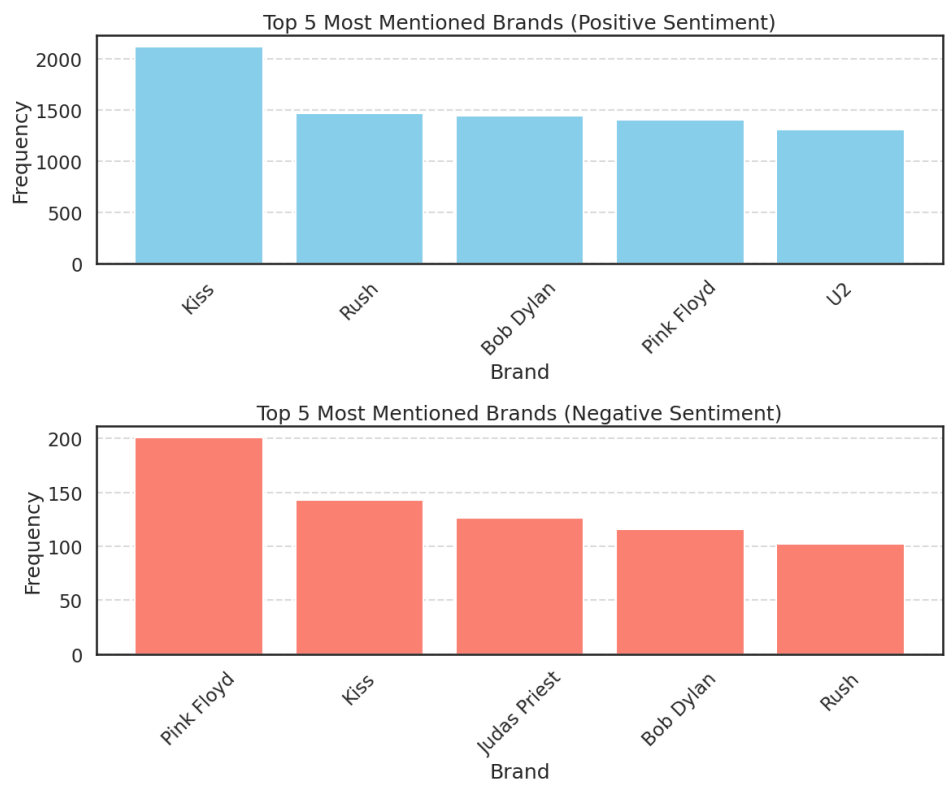


The analysis of network graphs from positive and negative reviews in the CDs and Vinyl category offers crucial insights into consumer sentiments. The positive reviews highlight words like "great," "album," and "songs" at the center, indicating strong satisfaction with the quality and enjoyment of music albums, particularly in the classic rock genre which is valued for its timeless appeal and standout tracks. Phrases like "highly recommended" and "best" suggest that consumers are not only pleased but are likely to recommend these albums.

Conversely, the negative reviews center around terms like "album" and "great song," yet are associated with words like "waste" and "money," pointing to disappointment and unmet expectations. Although some tracks are appreciated, the overall album experience often falls short, especially noted in mentions of specific bands like "Iron Maiden" and "Judas Priest" which indicates

particular dissatisfaction with these releases, potentially due to production quality or content. This dichotomy in sentiment provides valuable feedback for producers and marketers, highlighting the critical areas for improvement in product offerings and consumer satisfaction strategies. The network graphs thus serve as a strategic tool for aligning products and marketing efforts with consumer preferences and expectations in the evolving music industry.

Top 5 Most Mentioned Brands (Positive & Negative)



3.3 How do album features and metadata correlate with consumer purchasing behaviors and review ratings within specific music genres?

Correlation Matrix of Price, Review Year, Ratings, and Sentiment Score



### **Genre-Specific Correlations:**

The correlations between price, overall ratings, and sentiment scores within specific music genres like Alternative Rock, Classic Rock, and Rock are also revealing:

#### **Price and Overall Ratings:**

Alternative Rock: A very slight negative correlation (-0.004802), suggesting no meaningful relationship between price and ratings.

Classic Rock: A weak positive correlation (0.027125), suggesting that higher prices might slightly correlate with better ratings, but the relationship is very weak.

Rock: A slight negative correlation (-0.013707), similarly indicating no significant relationship.

#### **Price and Sentiment Scores:**

Alternative Rock: A weak positive correlation (0.021902), suggesting a very weak trend where higher prices might correlate with slightly more positive sentiment.

Classic Rock: Virtually no relationship (-0.003373).

Rock: A weak negative correlation (-0.021732), suggesting that higher prices could slightly correlate with more negative sentiment, though the relationship is weak.

The correlation matrix and genre-specific data from the CDs and Vinyl dataset reveal key insights into the relationship between album features, metadata, and consumer behavior. Notably, there is a minimal correlation between price and overall ratings across genres, indicating that pricing might not significantly influence consumer satisfaction or purchasing decisions. However, in genres like Classic Rock, a slight positive correlation suggests that higher prices could be linked to higher ratings, possibly due to perceived quality or the collectible nature of these items. Furthermore, the stable correlations between review year, ratings, and sentiment scores suggest that consumer perceptions remain consistent over time, with factors like genre or artist popularity playing a more significant role than temporal trends in shaping consumer responses. The positive correlation between overall ratings and sentiment scores validates the use of sentiment analysis in assessing consumer satisfaction, highlighting that more favorable reviews typically carry a positive emotional tone. These findings suggest that marketing strategies should focus more on enhancing product features such as packaging, exclusive content, and overall quality—especially in sensitive genres—rather than on competitive pricing. This approach is supported by the consistent consumer preferences over time, indicating opportunities for long-term brand building and loyalty programs

## 4. Conclusion

Based on the "CDs and Vinyl" dataset from Amazon, our analysis reveals a clear preference for Classic Rock, Rock, and Alternative Rock, which dominate in review volume and display strong positive sentiments, reflecting their lasting appeal and cultural significance. The sentiment analysis of review texts shows that positive reviews frequently contain words like "great," "love," and "best," which reflect the emotional impact and quality valued in these music purchases. Conversely, negative reviews often contain terms like "disappointing," indicating a gap between consumer expectations and actual experiences. This sentiment bifurcation provides a deeper understanding of consumer satisfaction and dissatisfaction drivers within the music industry.

Further, our correlation analyses between album features such as price and consumer ratings indicate that while pricing and the timing of releases have minimal impact on ratings, genre-specific variations exist. For example, a slight positive correlation in Classic Rock suggests a niche market where premium prices might reflect perceived collectability. The consistent preferences shown over time and across genres underline a stable interest in specific music types, allowing industry stakeholders to focus on enhancing product quality and emotional resonance rather than adjusting to fleeting market trends. This study not only underscores the cultural and emotional significance of music genres but also offers actionable insights for aligning product strategies with consumer expectations, fostering a strategic approach to product development and marketing in the music industry.

## 5. References

- Barbosa, L. N., Ribeiro, H., Tavares, F., & Santos, L. (2015). Predictive models for hospitality service ratings: Insights from sentiment analysis. *International Journal of Hospitality Management*, 49, 117-127.
- Budiharto, W., & Meiliana, M. (2018). Applications of sentiment analysis in the Indonesian elections. *Procedia Computer Science*, 135, 473-478.
- Guo, S. (2023). Vinyl records in the digital era: A renaissance of analog music. *Musicology Today*, 34(1), 88-102.
- Jain, A. (2013). Evaluating the entertainment industry through sentiment analysis: A predictive modeling approach. *Media Psychology*, 16(2), 207-224.
- Komara, E. (2015). Youth and vinyl: A new era of analog listeners. *Journal of Popular Music Studies*, 27(3), 325-345.
- López Medel, J. (2014). Visual artistry and symbolic expression in vinyl album covers. *Journal of Cultural Heritage*, 15(5), 532-539.
- Mittal, A., & Goel, A. (2012). Stock market prediction using sentiment analysis: A methodological improvement. *Journal of Financial Data Science*, 4(3), 54-65.
- Nguyen, T., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, 42(24), 9603-9611.
- Ramteke, J., Shah, S., Godsay, M., Shaikh, A., & Mishra, S. (2016). Election prediction via sentiment analysis of Twitter data. *International Journal of Computer Applications*, 139(11), 23-26.
- Tseng, F. S. C., Chou, A. Y. H., & Huang, J. H. (2017). Forecasting product pricing trends using sentiment analysis of online reviews and search engine query data. *Journal of Business Research*, 70, 94-101.
- Yadav, M., Krishna, A., & Chakraborty, T. (2019). Harnessing Twitter for forecasting the buyer's behavior in the e-commerce space. *Decision Support Systems*, 114, 26-37.