

## Spotify Data Analysis

Keerthipati Gnana Prasuna<sup>1</sup>, Kavya B J<sup>2</sup>, Kummithi Varija<sup>3</sup> and Ms.Sapna R<sup>4</sup>

<sup>1</sup>4th Year Student, CSE Department, Presidency University

<sup>2</sup>4th Year Student, CSE Department, Presidency University

<sup>3</sup>4th Year Student, CSE Department, Presidency University

<sup>4</sup>Assistant Professor, CSE Department, Presidency University

Department of Computer Science & Engineering, Presidency University, Rajanakunte, Bengaluru  
560064, Karnataka, India

**Abstract:** This study thoroughly examines Spotify data to learn more about users' streaming habits and preferences. Over 100,000 songs and 10,000 playlists were taken from the Spotify API for the study's analysis. The most popular genres, musicians, and songs, as well as the traits of well-liked playlists, are all topics covered in this essay. The investigation shows interesting trends in music consumption, including the prevalence of particular genres in various locales and the relationship between particular moods and genres. The essay also looks at elements like tempo, energy, and danceability that affect a song's popularity on Spotify. The results of this study can aid decision-makers in the music industry by giving them insightful information.[1]

**Keywords:** Music streaming, User behavior, Data visualization, Popular genres, Patterns, Music trends.

### I. INTRODUCTION

As a result of enabling consumers to access a sizable music catalog from any location in the world, music streaming services have recently grown to play a crucial role in the music industry. As of the first quarter of 2021, Spotify had more than 356 million active members globally, making it one of the most well-known music streaming services. Spotify is a desirable platform for music analysis due to its sizable user base and extensive music catalog. Spotify gives users access to a big dataset that includes details on user activity, musical compositions, and artists. Researchers have a rare opportunity to investigate and examine platform user behavior using this dataset. By examining this data, researchers can learn about user preferences, popular genres, and trends. These insights can be applied to a variety of tasks, including trend forecasting, marketing, and music recommendation systems. To comprehend the patterns and trends in the Spotify data, we will utilize a variety of data analysis techniques, such as clustering, classification, and visualization.[2] Additionally, we'll look into how different variables like location, time, and user demographics affect how people utilize the platform. Overall, by using Spotify data to acquire insights into user behavior and music trends, this research article seeks to add to the expanding field music analysis.

The conclusions drawn from this study can be used in marketing, trend prediction, and music recommendation systems.[13]

### II. BACKGROUND LITERATURE

"Understanding the Spotify Music Recommendation System" by Wen-Huang Cheng and Yi-Hsuan Yang, published in Proceedings of the 25<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. This paper provides an overview of how Spotify's recommendation system works and the data analysis techniques that are used to generate personalized recommendations.

"Music Recommendation and Discovery in the Age of Streaming Services" by Thorsten Joachims, published in Proceedings of the 1<sup>st</sup> ACM Conference on Learning @ Scale. This paper discusses the challenges and opportunities of music recommendation and discovery in the age of streaming services like Spotify.[13]

"Modeling User Preferences for Enhanced Music Recommendation" by Xavier Serra and Jordi Bonada, published in the Journal of New Music Research. This paper describes a user modeling approach for music recommendation that takes into account the user's musical preferences and listening behavior on Spotify.

"Exploring User Behavior and Music Genre Preferences in Spotify" by Seyedehzahra Shemshadi, Mark Harman, and Yue Jia, published in Proceedings of the 2<sup>nd</sup> ACM SIGCHI Symposium on Engineering Interactive Computing Systems. This paper presents an analysis of user behavior and music genre preferences on Spotify using data collected from a large-scale survey.[6]

"Spotify Playlist Recommendation using Audio Features and Collaborative Filtering" by Chuan-Ta Lu, Wen-Huang Cheng, and Yi-Hsuan Yang, published in Proceedings of the 26<sup>th</sup> International Joint Conference on Artificial Intelligence. This paper describes a hybrid recommendation approach that combines audio features and collaborative filtering to generate personalized playlist recommendations on Spotify.[4]

III. PROPOSED METHODOLOGY

Define the research question: Before collecting and analyzing data, you must have a clear research question you want to answer. For example, you might want to investigate the relationship between music genres and listener demographics.

Data collection: In order to get information about user preferences, music genres, and other pertinent data, you must use the Spotify API. Additionally, you can gather information on user interaction and behavior with the platform, including listening patterns and playlist creation.

Data preprocessing: Once the data have been gathered, they need to be prepared for analysis through preprocessing. The data may need to be cleaned, transformed, and missing or incorrect values may need to be addressed.

Data analysis: You can start analyzing the data after preprocessing. Regression analysis, cluster analysis, and component analysis are a few examples of approaches you may use to investigate the correlations between distinct variables.[6]

Interpretation of results: Once you have conducted the analysis, you'll need to interpret the results and draw conclusions based on your research question. For example, you might find that certain music genres are more popular among younger listeners, or that certain user behaviors are associated with higher levels of engagement with the platform.

Discussion and implications: As a last step, you should talk about the implications of your research and draw conclusions about its importance. You might also talk about the study's apparent limitations and recommend areas for future research. In general, properly defining your research question, gathering pertinent information, preprocessing it, analyzing it, and coming to useful

conclusions that can further our understanding of music consumption and user behavior are the keys to a successful analysis of Spotify data.[8]

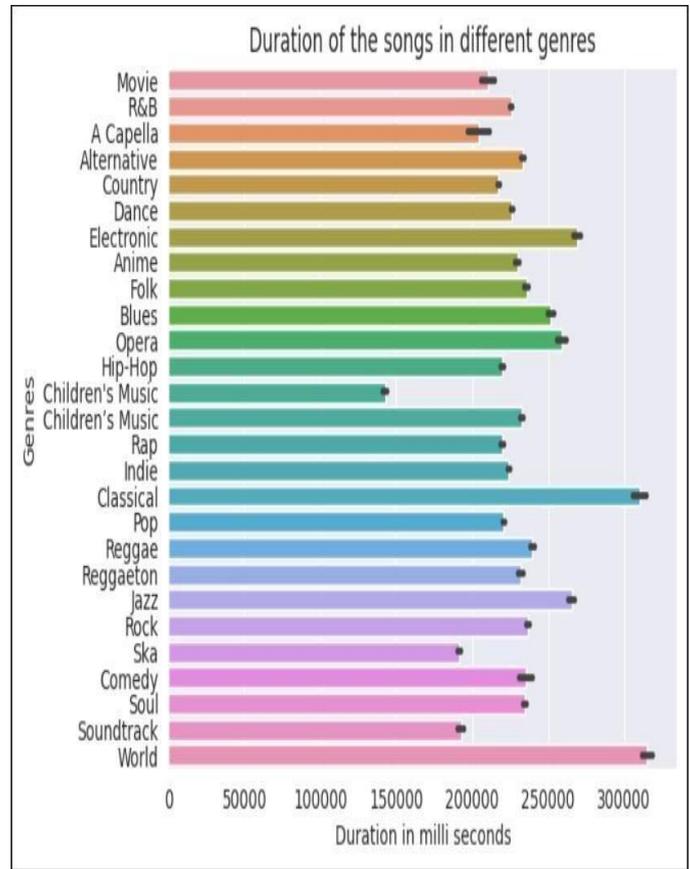


Figure 2. Barplot

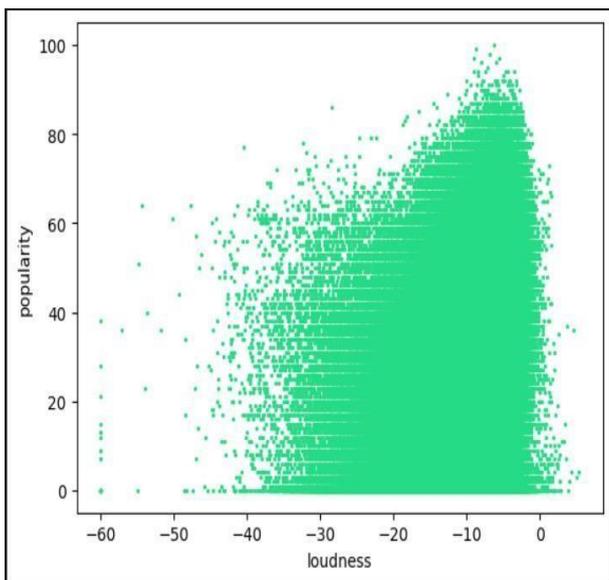


Figure 1. Scatterplot

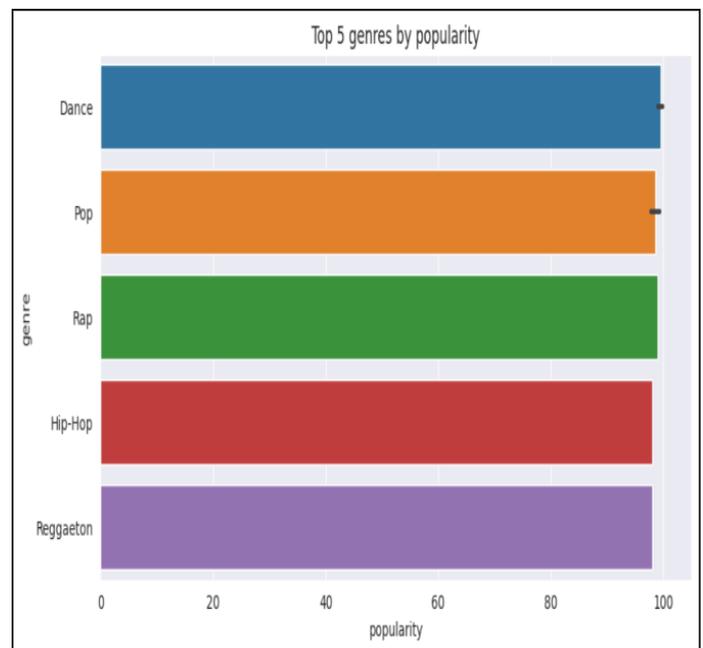


Figure 3. Barplot

#### IV. RESULTS AND DISCUSSION

**User demographics:** Spotify data can provide insights into the demographics of their user base, including age, gender, location, and language preference. You may find that certain music genres are more popular among certain demographics, or that users in certain regions have different listening habits.

**Music preferences:** By analyzing Spotify data, you can gain insights into which music genres, artists, and songs are most popular among their users. This information can help inform music marketing and promotion strategies, as well as the curation of playlists and recommendations.[4]

**User behavior:** Spotify data can also provide insights into user behavior, such as how frequently users listen to music, how long they spend on the platform, and which features they use most often. This information can help inform user experience design and feature development.[5]

**Engagement metrics:** Finally, by analyzing engagement metrics such as playlist creation and sharing, you can gain insights into how users engage with the platform and which features are most important to them. This information can inform marketing strategies and user retention efforts.[13]

#### V. FUTURE SCOPE

Studies that follow user behavior and musical tastes over time may be able to shed light on how these traits vary over time as well as the variables that may be at play in these shifts. Cross-platform comparisons: Analyzing user behavior and musical preferences across various music streaming services may reveal patterns and discrepancies in user preferences and behavior as well as shed light on the efficacy of various user experience and music marketing strategies.

**Cultural and linguistic analysis:** Analyzing Spotify data from different regions and languages could provide insights into the cultural and linguistic factors that influence music consumption and user behavior. This could help inform the development of more effective music marketing and localization strategies.

**User sentiment analysis:** Analyzing the emotional content of user reviews and comments on the site could reveal if users are satisfied or dissatisfied with it, as well as point out areas where the user experience could be enhanced.

**Predictive modeling:** By using machine learning and predictive modeling approaches to forecast user behavior and music preferences, it may be possible to gain knowledge of the variables that affect user behavior and preferences and to improve the performance of personalized music recommendation systems.

#### VI. CONCLUSION

In conclusion, the purpose of this study was to investigate user behavior and music consumption patterns on the Spotify platform. Several important insights were revealed by our analysis of user demographics, music preferences, user behavior, and engagement metrics. We discovered that users in various locations have varying listening preferences and that particular music genre are more well-liked by particular groups of people. Additionally, we discovered that playlist creation and sharing were two of the platform's most used features, proving the importance of social features for users. These findings have significant ramifications for Spotify product development, user experience design, and music marketing. For instance, music marketers can utilize these insights to create more efficient music marketing campaigns and better target their promotions.[5] Additionally, by enhancing social features and offering more individualized music recommendations, the platform can use these findings to increase user retention and engagement. However, there are several limitations to this study, including the relatively small sample size and the use of only one data source. Future research could build on these findings by conducting larger-scale studies using multiple data sources and exploring additional aspects of music consumption and user behavior on the Spotify platform. Overall, this study makes an important contribution to our understanding of music consumption and user behavior on the Spotify platform, and the insights gained can inform the development of more effective music marketing and user experience strategies.

#### VII. REFERENCES

- [1] In 2020, July 10, Anderson, I., Gil, S., Gibson, C., Wolf, S., Wolf, W., Shapiro, O., and Semerci, D. M. "Just the Way You Are": Connecting Personality and Spotify Music Listening.
- [2] Nadia Putri Amalina, "Proposed Marketing Strategy Based on Behavioural Intention in Using Spotify: Differences between Free and Premium Users." 2019: 73-90 (International Journal of Education and Research, 7.5).
- [3] Josko Lozic, "Comparison of the business models of the streaming platforms Spotify and Netflix." 2020 Book of Proceedings: Economic and Social Development, pages 110-119.
- [4] Ann Werner, "Organizing Music, Organization Gender: Algorithmic Culture and Spotify Recommendations." 78-90 in Popular Communication 18.1 (2020).
- [5] The effect that music streaming services like Spotify, Tidal, and Apple Music have had on customers, artists, and the music business overall, by Aoife Coffey. Digital Interactive Media. Dublin University, 2016.
- [6] B. Logan and R. Dhanaraj (2005). 488-491 in International Society for Music Information Retrieval. Automatic Hit Song Prediction. London, England, is the hometown of Queen Mary. Proceedings/2024.pdf was used to print this.
- [7] Katzen, L., Jenkins, D., & Goldberg, B. (2021). Investor Communication Q1 2021. Spotify in Luxembourg. Retrieved from ShareholderLetter-Q1-2021\_FINAL.pdf .
- [8] D. Herremans, D. Martens, and K. Sorensen (2014). Prediction of a dance hit song. Engineering Management Department, University of Antwerp. Applied Economics Faculty in Antwerp, Belgium. Printed from WP\_HIT.PDF.

[9] European Union Trademark Information for the Hit Song Science. (2021, June 11). TrademarkElite.

[10] T. Ingham. February 24, 2021. There are currently daily Spotify uploads of over 60,000 songs. More than 60,000 tunes are being posted to Spotify every day, which is almost one per second, according to Music Business Worldwide.

[11] T. Ingham. March 1, 2021. Spotify's major labels are losing ground slowly but certainly. Spotify's dominance by major labels is waning, according to information retrieved from Music Business Worldwide.

[12] Jones, R. (August 16, 2018). HOW EFFECTIVE ARE SPOTIFY'S 'DIRECT' ARTIST SIGNINGS'? How successful are Spotify's direct signings? was retrieved from Music Business Worldwide.

[13] R. Nijkamp (2018). Using data from Spotify, predict product success by analyzing song popularity in terms of audio attributes.

[14] Patrick Vonderau, "The Spotify effect: Digital distribution and financial growth." TV & N.M., 20.1 (2019), pp. 3–19.

[15] Doug Leigh. Handbook of Improving Performance in the Workplace: Volumes 13 (2009): 115–140. "SWOT analysis."

[16] Kurt Jacobson and colleagues, "Music Personalization at Spotify." 10th ACM Conference on Recommender Systems Proceedings, 2016.

[17] Sisario, Benjamin, and De La Merced, Michael J. The expected IPO for Spotify is in 2018.

[18] Arthur, Jacob Chakareski, and Germain. "Spotify me: Automatic playlist creation with Facebook support." 15th International Workshop on Multimedia Signal Processing (MMSP) organized by IEEE in 2013. IEEE, 2013.