Proposal: Understanding the Impact of Spatial Audio on User Behavior

Evan Gray

1) Introduction

The primary issue this project attempts to solve is the understanding of how spatial audio affects the user experience for listeners of music. Spatial audio is an upcoming technology with potential applications to music, film and virtual and augmented reality. The goal of which is to provide users with a more immersive experience. While there is significant research on the benefits of spatial audio, it tends to focus on either the technical side of making it work or the user experience of using spatial audio. [1,2,3] Studies conducted by companies such as Dolby Labs have indicated that users have a strong preference for sound quality when it comes to choosing a streaming service. [4] Additionally, of the 6000 respondents to Qualcomm's 2022 State of the Sound Report survey, 41% of users indicated an interest in spending more for spatial audio enabled devices. [5] There appears to be a major gap in research relating to how spatial audio influences a user's interaction with music. This research aims to use engagement metrics like listening time, skip rates and user satisfaction to quantitatively compare a user's experience with spatial audio versus stereo audio.

In recent years, streaming services have become a major avenue for individuals to interact with music and Spotify has emerged as a dominant force in the industry. I see Spotify as the primary sponsor of this project as it has over 31% of the market share in the music streaming services industry. [6] Direct sponsorship by Spotify would provide me with streamlined access to user data and would have a vested interest in understanding user engagement. The platform has over 640 million users globally, over 250 million of which are paid subscribers. [7] By leveraging their extensive user base and existing infrastructure, Spotify can directly benefit from actionable insights on how spatial audio influences user behavior.

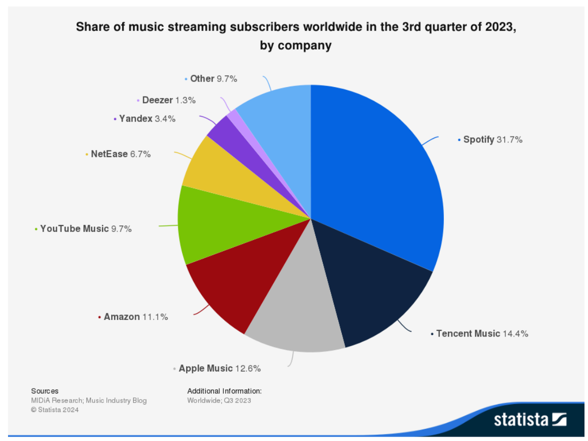

Figure 1: Market share of various music streaming services. [6]

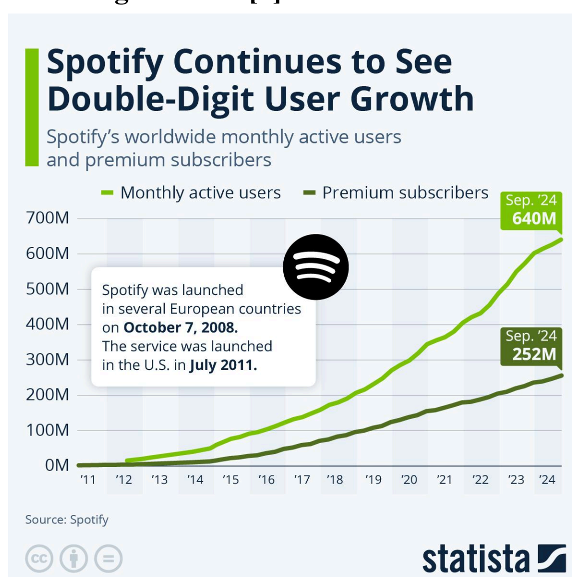

Figure 2: Growth of Spotify user base since 2011. [8]

Simultaneously with the rise of streaming services, improving sound quality and building additional features have been at the forefront of consumer audio electronics companies. Spatial audio sound formats such as Dolby Atmos and Sony 360 Reality Audio are becoming increasingly integrated in new products. [9] Determining the influence of spatial audio can help guide content creation, marketing strategies and strategic business partnerships. If spatial audio is found to significantly enhance engagement, it could justify further investments in the technology for a consumer electronics company and encourage spatial audio formats for content creation. All of this supports the retention of users and differentiation of products from the market. Further, streaming companies and consumer audio electronics companies could form partnerships to improve support for spatial audio. As with any industry, the state of the art must constantly be pushed to remain relevant. This unique potential relationship between streaming platforms and electronics manufacturers highlights the broader audio industry implications of the research.

I am personally interested in this project because listening to music is a fundamental part of my life. It helps me process emotions and bolsters the enjoyment of everyday activities. An upbeat song can set my mood for the day while a softer song can help me relax before bed. This project joins my love of music with my passion for data analysis. I am deeply interested in understanding a consumer base and enjoy working with cross-functional teams to determine the best features and use cases for data models. I have a quantitative and business-oriented mindset and I thrive on considering both the short-term and cascading effects of decisions. This project is a way to shape the future of user experiences through data-driven insights.

2) Motivation

Current solutions regarding the use of spatial audio tend to focus on improving the technology rather than analyzing user engagement with the format. Papers such as Potter et al. (2022) [1] and Cobos et al. (2022) [2] focus on the impact of spatial audio on virtual reality immersion and spatial audio methods without much discussion of how spatial audio actually influences the behavior of its users. Other studies have focused on the health benefits of spatial audio such as Johnston et al. (2022) [3] which examines how spatial audio may reduce auditory hypersensitivity in individuals with autism. Similarly, while Moiragias and Mourjopolous (2023) [10] look at how audio format contributes to emotional response and overall listening experience, participants listened to a predefined set of audio clips rather than engaged with their preferred music types. Although this is effective for understanding the desired relationships, the approach fails to account for individual preferences and listening behavior. All of these studies, while important, ultimately exist in a controlled, academic environment. They contribute to an empirical understanding of how spatial audio interacts with a person but they fail to examine the converse–how an individual interacts with spatial audio.

The proposed solution stands apart from existing approaches by tailoring its analysis of spatial audio usage to individual music preferences and behaviors. This project would focus specifically on analyzing who is using spatial audio and how they are using it. Likewise, it would include a comparative analysis to understand how engagement patterns differ between comparable segments of users and non-users of spatial audio. Spotify already collects and segments its user base for use in its advertising business and for artists to understand their listeners. [11,12] Unlike the current solutions that primarily focus on advancing spatial audio technology, this approach dives deeper into understanding the actual usage patterns of listeners and offers a real-time view into constantly shifting user behavior. By leveraging data collected from listening sessions, this analysis aims to go beyond lab settings and capture the nuanced ways individuals interact with spatial audio in their daily life. In doing so, this research fills a critical gap in our understanding of spatial audio adoption and usage.

Solving this issue would offer immediate benefits beyond those for the primary beneficiary. The most apparent would be the allocation of resources toward advancing spatial audio research which could lead to improved audio quality and better overall user experiences. Increased focus on spatial audio would also generate heightened public interest in new audio innovation. Enhanced enjoyment of music and media as a result of broad spatial audio usage could inspire innovation in related fields such as augmented and virtual reality. This could lead to a widespread adoption of spatial audio which would have a cascading effect on the overall user experience across various digital media platforms.

In the long run, solving this issue could significantly enhance music listening experiences, which would allow more people to enjoy immersive audio that deepens their connection to music. Given the positive effects of music on mental health, improved audio experiences could contribute to overall well-being and stress relief. [19] Additionally, a better understanding of spatial audio's benefits may spark further research into its potential and encourage advancements in audio technology as a whole. In the long run, increased competition for the best spatial audio experience could drive down costs and make spatial audio accessible to a wider audience. As a result of decreased costs, spatial audio could have a global effect on entertainment, mental health and digital engagement.

If spatial audio is found to drive user engagement, music streaming platforms and artists could be key beneficiaries of this research. This could lead to a short-term rise in the cost of spatial audio headphones as manufacturers respond to demand and advancements in hardware. Greater insight into user behavior would allow companies like Spotify to enhance their spatial audio offerings and drive them to adopt spatial audio formats as a listening option. This increased complexity and costs, however, may increase the barrier to entry for music engineers and independent artists. [20] This may lead to challenges for those unable to afford the resources to produce high-quality spatial audio and negatively impact the diversity of offerings on streaming platforms. Ultimately, while there is potential for a more immersive audio experience, the solution may also influence accessibility in both hardware and content production.

3) Data Requirements

The scale and complexity of this project necessitate big data infrastructure. Over 11 months in 2023, Taylor Swift, the top streamed artist of the year had 26.1 billion streams which would account for an average of almost 80 million streams daily. [13] This example illustrates the massive amounts of listening history data that Spotify has access to and processes daily. Moreover, the real-time nature of data collection and analysis adds an additional layer of complexity. This requires scalable parallelization to efficiently filter, process and model the data. Without big data techniques such as distributed computing and data lakes, analyzing these datasets would be computationally infeasible. By leveraging tools such as Apache Spark, this project can handle the high-velocity, high-volume nature of the data while extracting meaningful insights.

A foundational requirement to this solution is access to Spotify's music streaming session data. [14] This data includes but is not limited to metrics such as skips, session length and the reason a track started (e.g., user selection, previous track end). Each session also includes track metadata which has song-defining features such as popularity, tempo and danceability. In order to approach the spatial audio problem, this data must also include information regarding the device that played each track and whether they had a spatial audio feature turned on. This information would come from the app user listening logs which can be accessed across platforms such as Apple and Android devices. [15,16] To process and store this data, a platform such as Snowflake would be necessary to handle the large volumes of data efficiently. Given that the data may come from multiple sources, Snowflake is an ideal choice as it simplifies the implementation of extract, load and transform (ETL) pipelines. [17] Computing power from platforms such as Google Cloud Engine or Databricks will support Apache Spark for iterative, interactive analysis. [18] This will enable machine learning approaches and exploratory data analysis at scale. Spark is also of interest as it supports streaming data which could simplify the process of real-time analysis. These resources must scale with data volume so cloud solutions are a good choice to accommodate massive computation requirements. Beyond technological resources, this project may require access to in-house legal counsel as some data may be sensitive or subject to local privacy restrictions. As a result, the research may need to be limited to areas where such access is possible.

These resources were identified by looking at industry standards for companies that use big data approaches to business intelligence. Snowflake was chosen specifically for its relative ease of ETL pipeline implementation. Databricks was chosen as computation power can be scaled with actual usage. Additionally, Databricks clusters can be configured with Spark and MLFlow, a model tracking tool. Both tools scale well and can contribute to an understanding of user behavior shifts over time.

4) Technical Details

This project will primarily leverage Spotify's existing infrastructure and Apple's spatial audio APIs to collect data, and it will use frameworks such as Apache Spark to process and analyze the data.

The primary dataset for this study will consist of listening session data that tracks how an individual interacts with the service. Spotify already collects this data as evidenced by Brost, Mehrotra and Jehan (2019). [14] Key features to be built and used include session length, the number of tracks in each session and the number of skips in each session. Session length will be the primary feature of interest as it provides the most apparent engagement metric for the user. The number of tracks may serve as a secondary engagement metric, while the number of skips might have some implications for the behavior of the user. Further, the listening session data includes a track ID which can be used to gather metadata about each track. Such metadata can be used to paint a picture of each listening session which will generate insights about the types of individuals actually using spatial audio. This dataset will be complemented by an integration of Apple's Spatial Audio Usage API. [15] While this API would limit the dataset to Apple Users, this subset is significant, as Apple headphones account for approximately 35% of the U.S. headphone usage according to a 2024 study by Statista. [21] Additionally, the collection of such data may require an update to Spotify's privacy policy to secure user consent.

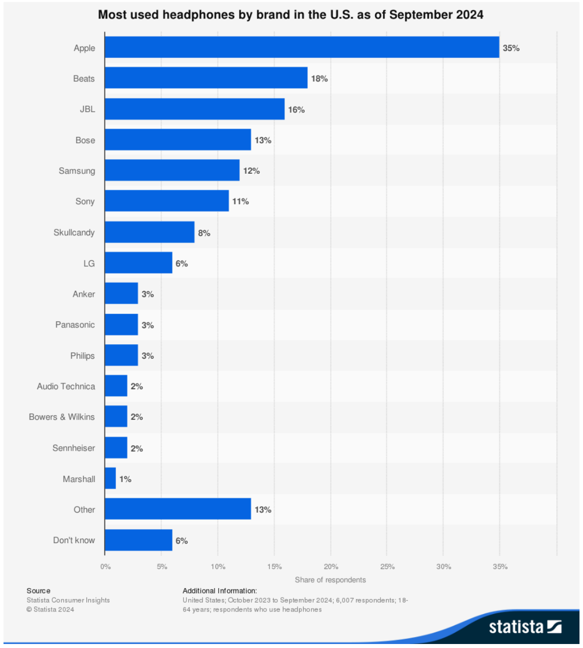

Figure 1: Most used headphones by brand in the U.S. as of September 2024. [21]

The analysis will focus on two primary comparisons using statistical tests such as T-Tests and ANOVA. First, a general comparison will be made between users who engage with spatial audio and those who do not. Secondly, a more controlled comparison will examine users with access to spatial audio and compare the behavior of those who choose to use the feature and those who choose not to. Each of these would involve in-depth segments of the users by taste to understand how spatial audio differs according to genre. To identify these segments, K-Means clustering will be employed. K-Means will partition the dataset into distinct user clusters based on features like listening habits, preferred genres and spatial audio usage. This clustering will enable personalized insights into how spatial audio affects different listener profilers. Further predictive models will be made with techniques such as Random Forest classifiers to predict whether a user is likely to engage with spatial audio based on their listening behavior and preferences.

This process will involve exploratory data analysis and machine learning techniques, both of which would require big data techniques to work with this data at scale. While Spotify uses its own data pipeline, there are a few key tools whose analogs are likely deployed by the company. I will use Snowflake and Databricks to create a scalable and efficient data pipeline that supports in-depth analysis and machine learning workflows.

The pipeline will use Snowflake, a cloud-based data warehouse as the centralized storage solution. Snowflake has a multi-cluster architecture which is ideal for the complex queries that may arise from the listening session and headphone data. [21] It has support for data formats such as JSON which may be required for working with the data from Apple's API. Snowflake also includes a feature called Snowpipe which enables continuous and automated data ingestion. This is essential as data from both sources may come in both batch and streaming formats at differing frequencies. The dataset would likely update with each listening session so for the purposes of this analysis, it is essential that data specifically from Apple devices are fed into the dataset. These two data sources will be integrated according to date, time and duration of session. The combined dataset is likely to be in the terabyte scale. The original Spotify listening session dataset that I used as a reference consisted of uniformly sampled sessions collected over an 8 week period in 2019. The dataset included 160 million listening sessions and comprised 350 gigabytes. [14] Accounting for the growth in users from 200 million to 640 million, I estimate that a year's worth of data will account for between 7 and 10 terabytes. Once the data is loaded, I will use Snowflake's SQL based transformation capabilities to clean, normalize and join the data. From there, I will have the requisite datasets for each of the comparisons.

For the analysis portion of the work, I will use Databricks. This platform integrates with Snowflake and allows me to process the data from the warehouse without costly duplication. [23] It has built-in distributed computing capabilities that can be scaled depending on the workload. These capabilities are integrated with Apache Spark which makes it ideal for building machine learning models. Spark is well suited for iterative methods so it is a good choice for segmenting the users based on taste with a K-Means clustering approach. I will store the cluster labels and centers in a Snowflake table which will be useful in the delivery of results.

Since this is primarily an analysis whose outcome is of interest to business oriented individuals, I will use Tableau to dashboard the results. Tableau has a direct integration with Snowflake so querying the results is straightforward and efficient. [24] I will build interactive dashboards to display the differences between usage patterns in reference to spatial audio. The audience segmentation will allow the user to examine usage differences across tastes. In addition to the interactive dashboard, I will deliver a comprehensive written analysis of areas where behavior differs significantly depending on spatial audio usage. Additional dashboards may indicate the trends of spatial audio usage during real-world events such as the release of a new premium headphone set.

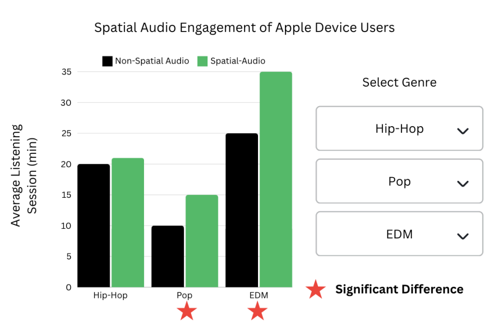

Figure 2: Example of dashboard with fabricated data.

The integration of these tools into Spotify's current toolset ensures minimal disruption. Snowflake's compatibility with existing SQL workflows and Databricks' seamless Spark integration make the proposed pipeline practical and relevant. Spotify's teams can use the findings to guide feature development, such as creating tailored playlists optimized for spatial audio or collaborating with hardware manufacturers like Apple or Bose to enhance compatibility. The beneficiary will use this dashboard to decide how to proceed with spatial audio integration. These tools would help Spotify's product and data teams make informed decisions about future feature development and user experience optimization. The teams could use these insights to better understand how spatial audio impacts user engagement and explore spatial audio features. For example, Spotify may decide to partner with headphone manufacturers to produce the best possible user experience across hardware and software.

5) Conclusion

Music streaming platforms have the goal of retaining and building their consumer base. With a deeper understanding of spatial audio usage, these companies can better refine their offerings and marketing to improve retention. They could use this information to inform content creators and artists as ways to improve streaming numbers. Consumer audio electronics companies could use this research to similarly market their products and drive sales of spatial audio-compatible products. I see this project as a win-win situation for both the sponsoring company and consumers of music. An understanding of engagement effects can have clear profit opportunities for the sponsor. Likewise, it would improve the reach of spatial audio which in turn would raise the quality of sound for the user. Music has been shown to have a positive effect on the well-being of individuals from improving social connections to supporting mental health. [3,10,19] With greater prevalence of spatial audio, these positive effects of music could have a wider scope.

The winners and losers can be identified based on factors like user engagement, cost implications, market competition and technological adoption. While streaming platforms that quickly adopt the technology may gain user retention, those slower to adopt may lose market share. Higher costs in the short run may negatively impact the users of spatial audio but they could ultimately improve user experience in the long run. This is also reflected in the competition of hardware manufacturers as supply in the market could drive costs down. Companies and creators willing to adopt spatial audio may emerge as winners in the short run while those constrained by entry costs could face challenges. If Spotify can understand the impacts of spatial audio on user behavior, it can remain ahead of rivals with a competitive advantage. This alignment between the analysis, insights and actionable strategies demonstrates the clear value for funding this project.

6) References

[1] Potter, T., Cvetković, Z., & De Sena, E. (2022). On the relative importance of visual and spatial audio rendering on vr immersion. *Frontiers in Signal Processing*, *2*. https://doi.org/10.3389/frsip.2022.904866

[2] Cobos, M., Ahrens, J., Kowalczyk, K., & Politis, A. (2022). An overview of machine learning and other data-based methods for spatial audio capture, processing, and reproduction. *EURASIP Journal on Audio, Speech, and Music Processing*, *2022*(1), 10. https://doi.org/10.1186/s13636-022-00242-x

[3] Johnston, D., Egermann, H., & Kearney, G. (2022). The use of binaural based spatial audio in the reduction of auditory hypersensitivity in autistic young people. *International Journal of Environmental Research and Public Health*, *19*(19), 12474. https://doi.org/10.3390/ijerph191912474

[4] *New survey from dolby reveals shifts in music listening behavior among adults in the u. S.* (2021, December 1). Dolby Newsroom. https://news.dolby.com/en-WW/206271-new-survey-from-dolby-reveals-shifts-in-music-listening-behavior-among-adults-in-the-u-s

[5] *The 2022 state of sound report*. (n.d.). Retrieved December 17, 2024, from https://assets.qualcomm.com/audio-state-of-sound-22-reg.html

[6] *Music streaming services subscribers market shares 2023*. (n.d.). Statista. Retrieved December 17, 2024, from https://www.statista.com/statistics/653926/music-streaming-service-subscriber-share/

[7] *Spotify—About spotify*. (n.d.). Spotify. Retrieved December 17, 2024, from https://newsroom.spotify.com/company-info/

[8] *Infographic: Spotify continues to see double-digit user growth*. (2024, December 5). Statista Daily Data. https://www.statista.com/chart/15697/spotify-user-growth

[9] *Dolby atmos: Spatial audio from the cinema to your headphones*. (n.d.). CNET. Retrieved December 17, 2024, from https://www.cnet.com/tech/home-entertainment/dolby-atmos-what-you-need-to-know-about-the-spatial-audio-format/

[10] Moiragias, G., & Mourjopoulos, J. (2023). A listener preference model for spatial sound reproduction, incorporating affective response. *PLOS ONE*, *18*(6), e0285135. https://doi.org/10.1371/journal.pone.0285135

[11] *The complete guide to audience segments in spotify for artists – spotify for artists*. (n.d.). Retrieved December 17, 2024, from https://artists.spotify.com/blog/introducing-new-audience-segmentation-on-spotify-for-artists

[12] Team, S. E. (n.d.). *Using demographic segmentation in marketing | spotify ads*. Retrieved December 17, 2024, from

https://ads.spotify.com/en-US/news-and-insights/demographic-segmentation-in-marketing/
[13] Stacy.Goldrick@groupsjr.com. (2023, November 29). *The top songs, artists, podcasts, and listening trends of 2023 revealed*. Spotify. https://newsroom.spotify.com/2023-11-29/top-songs-artists-podcasts-albums-trends-2023/
[14] Brost, B., Mehrotra, R., & Jehan, T. (2019). The music streaming sessions dataset. *The World Wide Web Conference*, 2594–2600. https://doi.org/10.1145/3308558.3313641
[15] *Spatial audio usage*. (n.d.). Apple Developer Documentation. Retrieved December 17, 2024, from https://docs.developer.apple.com/documentation/analytics-reports/spatial-audio-usage
[16] *Spatial audio and head tracking*. (n.d.). Android Open Source Project. Retrieved December 17, 2024, from https://source.android.com/docs/core/audio/spatial
[17] *What is an etl pipeline?* (n.d.). Snowflake. Retrieved December 17, 2024, from https://www.snowflake.com/guides/etl-pipeline/
[18] *Optimize the cluster utilization of Delta Live Tables pipelines with enhanced autoscaling*. (n.d.). Retrieved December 17, 2024, from https://docs.databricks.com
[19] Rebecchini, L. (2021). Music, mental health, and immunity. *Brain, Behavior, & Immunity - Health*, *18*, 100374. https://doi.org/10.1016/j.bbih.2021.100374
[20] Rys, D. (2024, February 8). Apple music's spatial audio royalty change raises indie label concerns over cost, artistry. *Billboard*. https://www.billboard.com/business/streaming/apple-music-spatial-audio-royalty-change-indie-labels-concerns-1235602510/
[21] *Most used headphones by brand in the U.S. 2024*. (n.d.). Statista. Retrieved December 17, 2024, from https://www.statista.com/forecasts/997153/most-used-headphones-by-brand-in-the-us
[22] *Multi-cluster warehouses | Snowflake Documentation*. (n.d.). Retrieved December 17, 2024, from https://docs.snowflake.com/en/user-guide/warehouses-multicluster
[23] *Read and write data from Snowflake*. (n.d.). Retrieved December 17, 2024, from https://docs.databricks.com
[24] *Snowflake and Tableau: Enhanced data storage and analysis.* (n.d.). Retrieved December 17, 2024, from https://www.tableau.com/solutions/snowflake
[25]