TECHNICAL REPORT ON SPOTIFY USER ENAGEMENT (2024)

Prepared by

Oludiran Ayoyinka Damilola

VEPH/20B/DA133

OUTLINE

1. Introduction
2. Story of Data
3. Data Splitting and Preprocessing
4. Pre-Analysis
5. In-Analysis
6. Post-Analysis and Insights
7. Data Visualizations & Charts
8. Recommendations and Observations
9. Conclusion
10. References & Appendices

1   INTRODUCTION

Spotify is a global digital streaming service that provides users with access to millions of songs, podcasts, and audiobooks. Launched in 2008, it uses advanced recommendation algorithms to personalize content based on user preferences, moods, and listening habits. With both free and premium subscription options, Spotify aims to enhance user engagement while generating revenue through ads and subscriptions. The platform seeks to ensure that whenever users visit, they feel their needs are met. Even for users who do not explicitly state their preferences, Spotify aims to recommend music and podcasts that align with their moods and desires at any time of the day.

This report presents an analysis of Spotify user engagement, focusing on preferred music genres, podcast genres, and subscription trends. The insights aim to enhance user experience and inform strategic decisions for content delivery.

Main Objective:

> The main objective of this project is to analyze user preferences in music and podcast genres to enhance recommendation accuracy to increase user satisfaction, engagement, and overall platform revenue.

Problem to be addressed

The global music streaming industry is highly competitive, with multiple platforms offering similar features and content. Streaming services like Spotify face challenges in attracting, engaging, and retaining users while ensuring long-term customer satisfaction. To maintain a competitive edge, it is crucial to understand user listening patterns, preferences, and behaviors.

This analysis aims to answer key questions such as:

- How do user demographics (age, gender, subscription type) impact listening behavior?

- What content (music vs. podcasts) do users prefer, and how does consumption frequency affect retention?
- What music genre do users prefer to listen to across different ages and genders?
- What format of podcast to users prefer to listen to across different ages, usage period and gen
- How can Spotify optimize its offerings to enhance user satisfaction and increase customer loyalty?

Key Datasets and Methodologies

The analysis is based on music streaming data, which includes user demographics, listening frequency, and subscription details. Microsoft Excel was used for data analysis, employing Pivot Tables to summarize user engagement patterns and Pivot Charts to visualize trends in music and podcast consumption. These tools helped uncover insights into user preferences, retention factors, and potential strategies to enhance customer satisfaction.

2. STORY OF DATA

i)   **Data Source**

The data used for this analysis was downloaded from the Kaggle repository, which was originally sourced from the company's internal database. It contains user information, listening habits, and subscription details. This dataset provides valuable insights into customer engagement, retention, and preferences.

ii)   **Data Collection**

The data was collected through first-time user registrations and daily listening activity on the platform. User details, such as demographics and subscription type, were recorded during registration. Listening patterns, including music and podcast consumption, were tracked automatically based on user interactions. This data was then compiled and stored in the company's internal database for analysis.

iii)   **Data Structure**

The dataset is organized in a tabular format, with each row representing an individual user's streaming activity. The columns include various variables such as User ID, Age, Gender, Subscription Plan, Music Listening Frequency, Podcast Listening Frequency, and Premium Plan Status. Additional fields capture listening duration, preferred genres, and engagement metrics. This structured format allows for easy analysis using Pivot Tables and Pivot Charts in Microsoft Excel.

iv)   **Importance Features and Their Significance**

a) **User Age Group:** Categorized as 12-20, 20-35, and 35-60 years. Helps understand music and podcast preferences across different age groups.

b) **User Gender:** Categorized as Male, Female, and Others in the dataset. Helps analyze differences in music and podcast preferences across genders.

c) **Preferred Listening Content (Music vs. Podcast):** Helps identify whether users engage more with music or podcasts. Useful for content recommendation strategies and targeted promotions

d) **Subscription Plan (Free vs. Premium):** Determines the ratio of free vs. paid users, influencing revenue strategies.Helps assess the effectiveness of premium features in driving upgrades.

e) **Premium Plan Preference:** Provides insights into the most popular premium subscription tiers.Helps optimize pricing and feature offerings to boost premium subscriptions.

f) **Leading Music Genre:** Identifies the most preferred music genres (e.g., Melody, Classical, Pop).Helps in curating personalized playlists and improving content recommendations.

g) **Music Influential Mood Preference:** Identifies the types of moods users prefer in their music (e.g., relaxation, motivation, sadness)Helps in curating personalized playlists and improving content recommendations.

h) **Peak Spotify Usage Duration:** Shows how long users have been active on the platform (e.g., more than 2 years, 6 months to 1 year). Helps in understanding retention trends and user loyalty.

i) **Top Trending Podcast Genres:** Identifies the most popular podcast genres (e.g., Comedy, Lifestyle & Health).Helps Spotify optimize content recommendations and support podcast creators.

j) **Most Preferred Podcast Format:** Identifies the types of podcast formats users prefer (e.g., storytelling, conversational).Helps Spotify tailor podcast recommendations and improve content diversity.

v) **Data Biases:** The dataset may not be fully representative of all Spotify users worldwide, as it is sourced from a specific internal database and Kaggle. Certain demographics (e.g., age groups, and geographic regions) may be underrepresented, which could skew insights.

3.   .**DATA SPLITTING AND PREPROCESSING**

i)   **Data Cleaning**

Before analyzing the Spotify user data, a data quality check was performed to ensure accuracy and reliability.This involved identifying missing values, duplicate entries, and inconsistencies using Microsoft Excel's built-in functions.

a)   **Handling Missing Values**

A prerequisite data quality check was conducted using Microsoft Excel's built-in functions to assess missing values. The check confirmed that there were no blank spaces in the dataset. This validation ensured data integrity and reliability before proceeding with the analysis. Missing values can distort statistical analysis and affect the accuracy of visualizations, leading to misleading conclusions. A dataset free of missing values ensures that calculations and comparisons remain consistent across all records. If missing values are present and not handled properly, they can introduce bias in model predictions or trend interpretations.

b)   **Removing Duplicates**

After checking for missing values, a duplicate records assessment was conducted using Microsoft Excel's built-in functions. The test confirmed that no duplicate entries existed in the dataset, indicating that data cleaning had occurred before the analysis. Duplicate records can distort statistical analyses and misrepresent trends. Removing duplicates ensures that each observation is unique, thereby maintaining the accuracy of the findings. Since no duplicates were found, further data cleaning was unnecessary. This validation reinforced the dataset's integrity, ensuring that the analysis would yield meaningful and reliable insights. By confirming the absence of duplicates,

the study upholds a high standard of data quality, ultimately leading to more accurate and actionable conclusions.

ii)   **Data Transformation:** No additional data transformations, such as normalization, standardization, or the creation of new variables, were performed before or after the analysis. The focus was primarily on data cleaning to ensure accuracy and reliability. Once the dataset was cleaned by handling missing values, removing duplicates, and correcting inconsistencies, the analysis was conducted directly using Microsoft Excel's Pivot Tables and Pivot Charts to explore user engagement, listening preferences, and subscription trends.

iii)   **Data Splitting:** In this study, age, gender, listening content, subscription plan, premium plan, music genre, music format, podcast genre, podcast format, and Spotify usage duration are the independent variables, while user engagement is the dependent variable.

iv)   **Industry Context:** The data pertains to the music streaming industry, particularly within the digital entertainment sector. The significance of this industry to the analysis lies in understanding user engagement, content preferences, and subscription trends to enhance customer retention and satisfaction. By examining music and podcast preferences, subscription plans, and listening behavior, Spotify can refine its recommendation algorithms, pricing strategies, and marketing campaigns to maintain a competitive edge. The insights gained from this analysis assist streaming platforms in personalizing user experiences, optimizing ad revenues, and increasing paid subscriptions.

v)   **Stakeholders**

a)   **Marketing Team** – Insights into user preferences and engagement patterns will help in crafting targeted campaigns, personalized recommendations, and promotions to attract and retain users.

b)   **Product Development Team** – Understanding popular music and podcast genres will guide the creation of new features, curated playlists, and improved recommendation algorithms.

c) **Senior Management & Strategy Teams** – Data on subscription trends and engagement levels will aid in decision-making related to pricing strategies, market expansion, and overall business growth.

d) **Content Creators & Artists** – Information on genre popularity and listening habits can help artists and podcasters tailor their content to meet audience demand.

e) **Customer Experience & Support Teams** – Identifying user pain points and engagement trends will assist in improving user satisfaction and resolving issues more effectively.

vi) **Value to the Industry**

Success in the music streaming industry is defined by user engagement, which is measured through listening duration, content interaction, and subscriber growth. This analysis aids decision-makers in optimizing recommendations, enhancing premium offerings, and refining marketing strategies to increase retention and revenue. Furthermore, insights empower content creators to customize offerings for a more personalized user experience. Ultimately, this fosters greater satisfaction, brand loyalty, and profitability in a competitive market.

4. **PRE-ANALYSIS**

i) **Identify Key Trend**

   a. **Most Preferred Listening Content by Age and Gender** – A trend may indicate that younger audiences prefer music, while older users engage more with podcasts.
   b. **Peak Usage Duration by Age and Gender** – Long-term Spotify users may exhibit higher engagement levels, whereas newer users contribute to a growing listener base.
   c. **Leading Music Genre by Popularity** – Certain age groups may favor specific music genres, with melody and classical potentially being more popular among older users.

d. **Top Trending Podcast Genres** – A pattern might suggest that entertainment-focused genres (e.g., Comedy) attract more listeners, while educational podcasts appeal to niche audiences.

e. **Most Preferred Subscription Plan by Age and Gender** – A trend may reveal that younger users opt for free plans, whereas working professionals are more likely to subscribe to premium plans.

f. **Most Preferred Premium Plan by Age and Gender** – Subscription choices may reflect financial capacity, with students favoring discounted plans and professionals preferring individual premium options.

g. **Top Music Influential Mood** – The influence of music on mood may show that relaxation and stress relief are primary drivers for music consumption.

h. **Most Preferred Podcast Format** – A possible trend indicates that storytelling podcasts may be more engaging than other formats, especially among younger listeners.

ii) **Potential Correlation**

a. **Listening Frequency vs. Subscription Plan** – A possible correlation may exist between higher listening frequency and users opting for premium plans due to ad-free experiences.

b. **Age Group vs. Preferred Content Type** – A trend might suggest that younger users engage more with music, while older users prefer podcasts.

c. **Podcast Genre vs. Subscription Plan** – Users who listen to educational or business podcasts may be more likely to subscribe to premium plans for exclusive content.

d. **Music Genre vs. Mood Influence** – A potential relationship may indicate that users who listen to melody or classical music primarily seek relaxation and stress relief.

e. **Study Hours vs. Listening Duration** – A trend might show that students who study longer prefer background music or instrumental genres to aid concentration.

f. **Subscription Plan vs. User Retention** – A correlation could suggest that premium subscribers exhibit higher long-term retention compared to free-tier users.

g. **Gender vs. Podcast Format Preference** – A noticeable trend may indicate that conversational podcasts are more popular among certain gender groups.

iii) **Initial Insights**

Before conducting a detailed statistical analysis, a preliminary review of the dataset reveals promising insights and areas that warrant further investigation:

a. Most Preferred Listening Content by Age and Gender
b. Peak Usage Duration by Age and Gender
c. Leading Music Genre by Popularity
d. Top Trending Podcast Genre
e. Most Preferred Subscription Plan by age and gender
f. Most Preferred Premium Plan by age and gender
g. Top Music Influential Mood
h. Most Preferred Podcast Format.

1. IN-ANALYSIS

Unconfirmed Insight

a) **Subscription Plan Preference**: The majority of users (424) showed a preference for the free (ad-supported) subscription plan over the premium (paid subscription) plan (96). However, this does not necessarily indicate dissatisfaction with premium plans, as factors such as affordability and perceived value might influence this preference.

b) **Premium Plan Selection**: A significant number of users did not express a clear preference for any premium plan, with 95 users showing interest in the individual plan. This suggests that while premium options are available, users might not see a compelling reason to upgrade or may prefer flexibility over commitment.

c) **Spotify Usage Duration**: The engagement analysis indicates that most users (169) have been using the streaming service for more than two years, while 91 users have used it for less than six months. This could imply that long-term users find sustained value in the service, whereas newer users may still be exploring their preferences.

d) **Listening Content Preference**: More than half (410) of the surveyed users showed a stronger interest in music over podcasts (110). While this suggests a dominant preference for music, it does

not rule out the potential growth of podcast listenership, as factors like content availability and personal habits may contribute to this trend.

e) **Music Genre Popularity**: Melody emerged as the leading genre with a frequency of 259, significantly ahead of classical music (87). This preference may be influenced by current trends, cultural influences, or the accessibility of certain genres on the platform.

f) **Trending Podcast Genres**: Users with no specific genre preference led with a frequency of 148, followed by Comedy, Lifestyle & Health, Health & Fitness, Sports, and Food & Cooking. The dominance of general preference suggests that users may prioritize content quality and engagement over strict genre classification.

g) **Influence of Music on Mood**: Relaxation and Stress Relief were the most frequently cited influential moods, with 195 users selecting them. This highlights the role of music in mental well-being, although individual mood preferences may vary based on external factors such as stress levels and daily routines.

h) **Podcast Format Preference**: Storytelling was identified as the most preferred podcast format. This suggests that users engage more with narrative-driven content, which may be due to its immersive nature and ability to hold attention. However, preferences may shift with emerging trends and content diversity.

ii)   **Recommendations**

a.   **Enhance User Acquisition Strategies for Premium Plans**

Since 81.5% of users prefer the free ad-supported plan, consider introducing targeted promotions or trials for premium plans to convert free users to paid subscribers. Highlight the benefits of premium features, such as ad-free listening, offline mode, and higher audio quality.

b.   **Optimise Premium Plan Offerings**

With 95 users showing interest in the individual premium plan, Spotify could explore bundling options or personalised offers to encourage uptake of other premium tiers, such as family or student plans.

c. Increase Engagement Among New Users

For the 91 users with usage under six months, consider onboarding initiatives such as personalised playlists, tutorials, or notifications to enhance engagement and retention.

d. Focus on Music Content

As music is significantly more popular than podcasts (410 vs. 110 users), continue curating music-based features, such as algorithm-driven playlists and artist-led content. Additionally, explore ways to boost podcast engagement, such as promoting popular podcast genres like Comedy and Health & Fitness.

e. Develop Genre-Specific Campaigns

With 'Melody' as the leading genre, consider playlist promotions, artist collaborations, or themed campaigns around this genre. Spotify could also explore opportunities to elevate less popular genres, like 'Classicals', through curated experiences or special features

f.   . Capitalise on Music for Relaxation and Stress Relief

Since 195 users prefer music for relaxation, Spotify can introduce mood-based playlists or partner with wellness brands for 'Stress Relief' campaigns, particularly around high-stress periods like exams or the end of the year.

g. Promote Storytelling Podcast Formats

Since storytelling is the preferred podcast format, Spotify could invest in exclusive storytelling podcasts or collaborate with influencers and content creators to produce compelling narratives.

h. Personalized User Experience with Data-Driven Insights

Utilise the insights on user preferences to create personalised recommendations, push notifications, or playlists tailored to individual listening habits.

iii) **Analysis Used in Excel**

a) **Pivot Tables**: Used to summarize user preferences across different categories, such as subscription plans, listening duration, preferred content, and music genre popularity. This allowed for efficient data organization and identification of key trends.

b) **Pivot Charts**: Created to visually represent relationships between variables, making it easier to interpret insights, such as the correlation between age and preferred subscription plans or the influence of mood on music choices.

c) **Conditional Formatting**: Applied color scales to highlight significant patterns, such as the most frequently chosen music genre, the top-ranking podcast categories, and high engagement levels based on listening duration.

2. POST ANALYSIS AND INISGHT

i) Key Findings

a. **Subscription Plan Preference:** 82% of users prefer the free, ad-supported plan, with the 20-35 age group and females showing the highest preference. This result is similar among long-term users (2+ years).

b. **Premium Plan Selection:** No strong preference for specific premium plans, but the individual plan had the highest selection. Females and the 20-35 age group had a majority with no clear preference.

c. **Engagement Duration:** Most users (169) have been on Spotify for over two years, with the 20-35 age group and females leading in long-term usage.

d. **Preferred Content-Type:** Music dominates over podcasts, with the 20-35 age group and females being the most engaged.It is preferred in the morning and at night, while podcasts gain some traction in the afternoon.

e. **Music Genre Popularity:** Melody is the most popular genre, particularly among the 20-35 age group and females. It dominates mornings and nights, while Classical is favored in the afternoon.

f. **Podcast Genre Trends:**148 users had no specific genre preference, but storytelling was the top choice among females and the 20-35 age group.vPodcast listening is mostly occasional, with storytelling as the most popular format.

g. **Music's Influence on Mood:** Relaxation and stress relief are the top reasons for music consumption.

h. **Podcast Format Preference:** Storytelling is the most preferred format, indicating demand for narrative-driven content.

ii) Comparison with Initial Findings

The analysis largely supports the initial expectations, with a few nuanced differences:

a) **Subscription Plan Preference** – The expectation that most users prefer the free, ad-supported plan was confirmed, with 81.5% favoring it. However, the strong preference among long-term users (2+ years) was unexpected, suggesting affordability is a key factor rather than dissatisfaction with premium options.

b) **Premium Plan Selection** – The assumption that users lack a strong preference for specific premium plans was validated. While the individual plan had the highest selection, most users showed no clear preference, reinforcing the idea that premium options may not offer compelling differentiation.

c) **Spotify Usage Duration** – The expectation that most users have been using Spotify for over two years was confirmed, with 169 users fitting this category. The 20-35 age group and females showed the highest engagement, suggesting sustained value for this demographic.

d) **Listening Content Preference** – The hypothesis that music is more popular than podcasts was supported, with 410 users favoring music over 110 for podcasts. However, the trend of higher podcast engagement in the afternoon was unexpected, indicating a potential opportunity for targeted content.

e) **Music Genre Popularity** – The prediction that Melody would be the most preferred genre was confirmed, with 259 users selecting it. However, its dominance at night over relaxing genres like Classical was an unexpected insight.

f) **Trending Podcast Genres** – The expectation that many users would have no specific podcast genre preference was validated (148 users). However, storytelling podcasts were more popular than anticipated, especially among females and the 20-35 age group.

g) **Influence of Music on Mood** – The assumption that relaxation and stress relief are major influences on music choices was confirmed, with 195 users selecting them, reinforcing the importance of mood-based playlist curation.

h) **Podcast Format Preference** – The hypothesis that storytelling would be the most preferred podcast format was validated. Its popularity highlights the demand for narrative-driven content, presenting an opportunity for podcast creators.
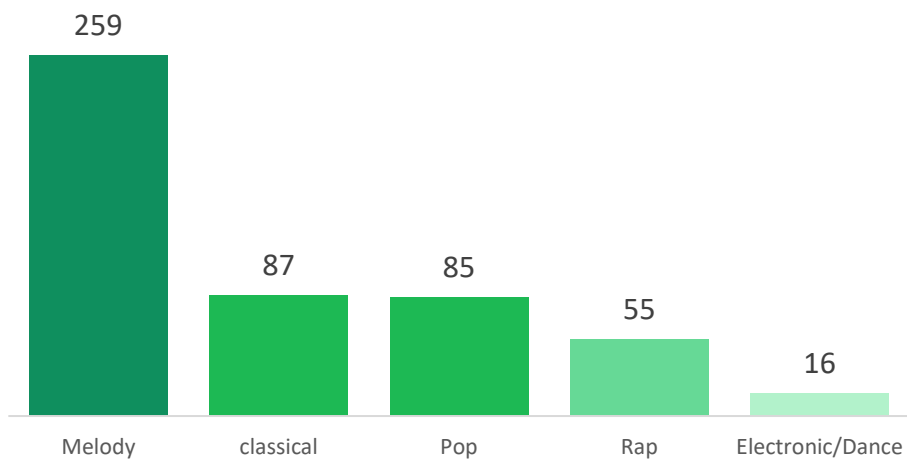
3. **DATA VISUALISATION & CHARTS**

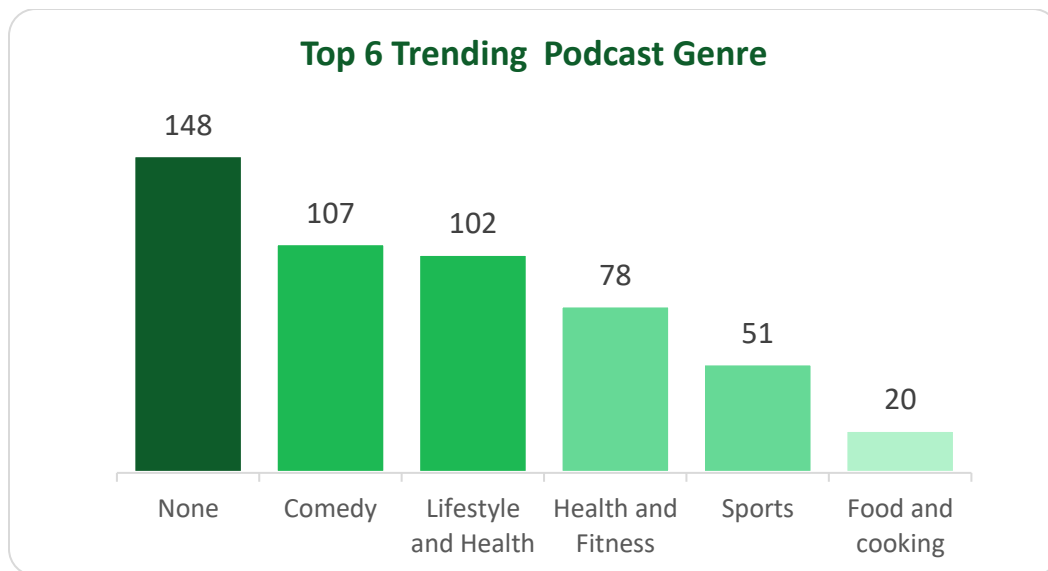**Most Preferred Listening Content**

21%

79%
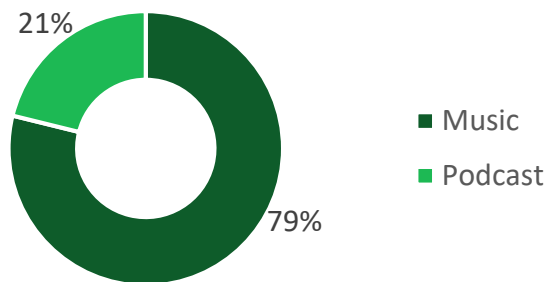
- Music
- Podcast

Statistics revealed that the majority (79%) of the users that made-up this data prefer listening to Music than Podcast.
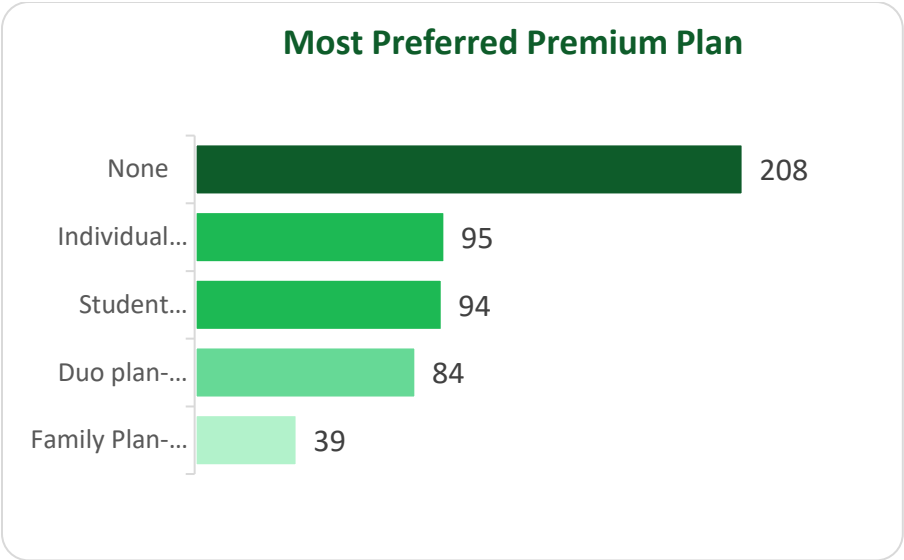
Leading Music Genre by Popularity

259

87

85

55

16

Melody   classical   Pop   Rap   Electronic/Dance

Among the 79% that prefers listening to Music, the majority (259) loves listening to Melody to other music genre.

Top 6 Trending Podcast Genre

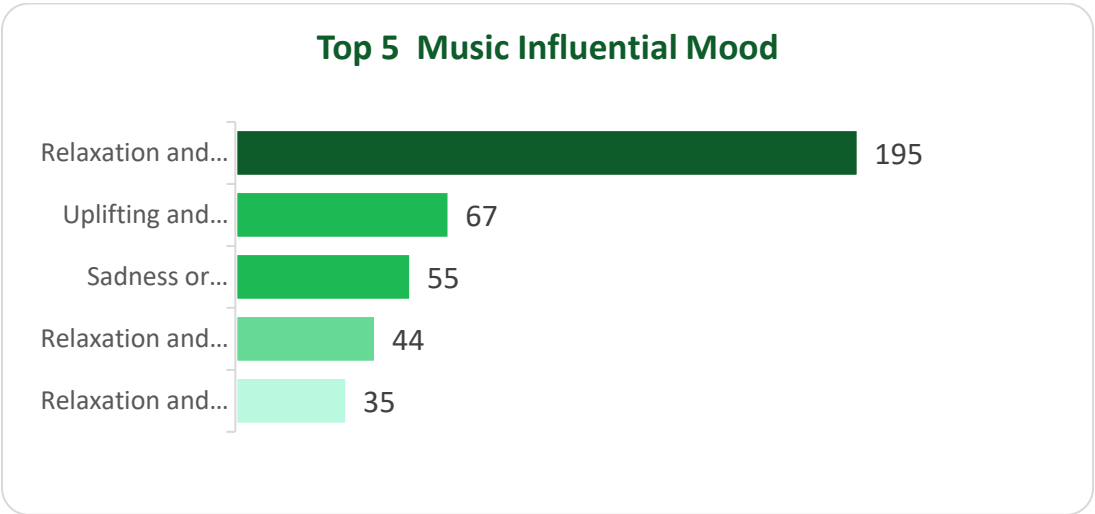|	
None	148
Comedy	107
Lifestyle and Health	102
Health and Fitness	78
Sports	51
Food and cooking	20

From the above chart, 148 users which is the majority revealed they have no specific preference for Podcast Genre, while 107 of them prefers listening to comedy, following (102) is lifestyle and health.
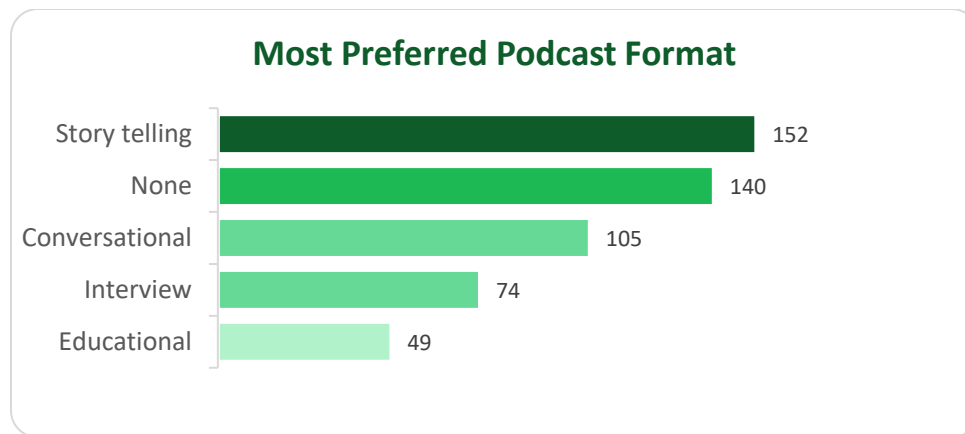
Most Preferred Subscription Plan

- Free (ad-supported): 82%
- Premium (paid subscription): 18%
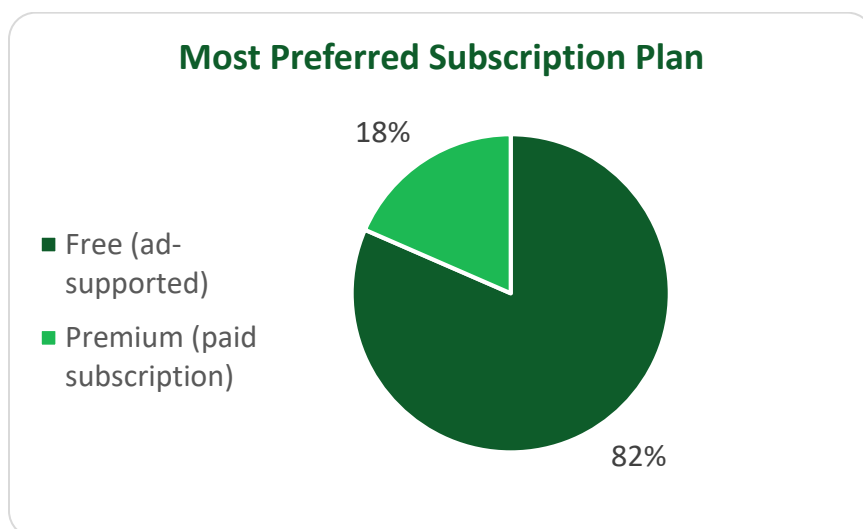
Most (82%) of the users showed interest in free (ad-supported) subscription to premium subscription (18%)

Most Preferred Premium Plan

| Plan	Count
None	208
Individual...	95
Student...	94
Duo plan-...	84
Family Plan-...	39

A large number (208)of respondents stated they had no preference for premium plans, this set of people must be those who opted for free subscriptions, the next highest were individual plans and student plan with 95 and 94 users showing interest respectively.

Top 5 Music Influential Mood

Mood	Count
Relaxation and...	195
Uplifting and...	67
Sadness or...	55
Relaxation and...	44
Relaxation and...	35

Among the lists of influential mood, 195 users which is the highest among the top 5 music influential mood, showed interest for relaxation and stress relief songs.

Most Preferred Podcast Format

Format	Value
Story telling	152
None	140
Conversational	105
Interview	74
Educational	49

From the chart above, storytelling (195) was revealed to be the most preferred podcast format among users that made up the data.
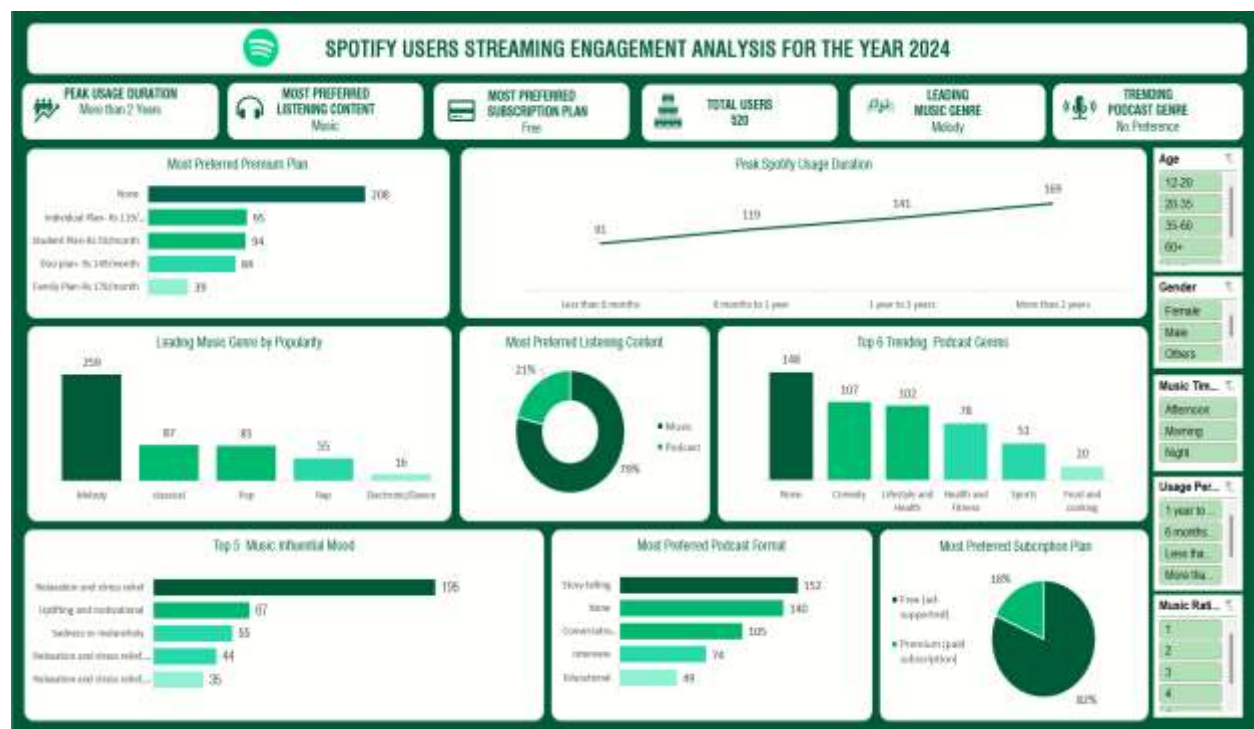
This dashboard comprises of all the charts used in the analysis of Spotify User Streaming Engagement for the Year 2024.

4. RECOMMENDATION AND OBSERVATION

a) **Subscription Plan Strategy** – Since affordability is a major factor in subscription choices, Spotify can introduce a low-cost premium tier with essential features to encourage upgrades. Additionally, offering long-term user discounts could help convert free users to paid subscribers.

b) **Premium Plan Enhancement** – The lack of strong preference for premium plans suggests the need for better differentiation. Spotify could introduce more customized premium features such as exclusive playlists, enhanced audio quality, or bundled deals (e.g., music + audiobooks).

c) **User Retention & Engagement** – With a high number of long-term users, Spotify should focus on retention strategies like loyalty rewards, exclusive content for long-time users, and AI-driven personalized recommendations to sustain engagement.

d) **Podcast Growth Opportunity** – Given the afternoon peak in podcast consumption, Spotify could promote midday podcast recommendations and create exclusive short-form content designed for break times (e.g., 10–15 min episodes).

e) **Music Genre Insights** – The strong preference for Melody at night suggests an opportunity to create personalized nighttime playlists. Spotify could also experiment with relaxing Melody sub-genres to attract users who might prefer calming music over traditional Classical.

f) **Podcast Genre Expansion** – The rising popularity of storytelling podcasts presents an opportunity to invest in original storytelling content. Spotify could collaborate with popular authors, scriptwriters, or influencers to create engaging, narrative-driven series.

g) **Mood-Based Playlists** – Since relaxation and stress relief are major reasons for music listening, Spotify should optimize its mood-based playlist curation and introduce AI-driven mood-matching recommendations for users based on listening patterns.

h) **Podcast Format Innovation** – With storytelling being the most preferred format, Spotify should experiment with interactive storytelling podcasts, where users can influence the storyline, making content more engaging and immersive.

CONCLUSION

The analysis confirms that most Spotify users prefer the free, ad-supported plan, with affordability being a key factor in their decision. Premium plan users showed no strong preference for specific options, indicating the need for better differentiation and added value. Long-term engagement remains high, particularly among the 20-35 age group and females, suggesting a loyal user base. Music remains the dominant content choice, but podcasts show growth potential, especially in the afternoon. Melody emerged as the most popular music genre, with a surprising dominance at night, while storytelling stood out as the preferred podcast format. The influence of music on relaxation and stress relief highlights the need for enhanced mood-based playlist curation. Overall, Spotify can leverage these insights to refine its subscription strategy, improve premium offerings, and create targeted content to maximize user engagement.

REFRENCE

The dataset was downloaded from Kaggle:

https://www.kaggle.com/datasets/meeraajayakumar/spotify-user-behavior-dataset/data