DAI-101 Assignment

Name - Krish Singla
Enroll. No. - 23114050
Branch - Computer Science and Engineering

Project Overview

This project performs Analysis on a Music Streaming dataset to explore relationships between categorical and numerical variables. Various visualization techniques are used to understand the impact of factors like age group,country etc on listening time, music platform, subscriptions etc. This is a part of the assignment given in the course DAI-101.

Flow of code

- Refer readme.md for that.
- To keep the report clean and concise I will include only final takeaways from data analysis.

What have I done in the project?

How have I leveraged data analytics to uncover user engagement patterns and drive strategic insights for business growth?

Feature analysis - Categorical or Numerical

- Categorical Columns: Country, Streaming Platform, Top Genre, Most Played Artist, Subscription Type, Listening Time (Morning/Afternoon/Night), Age Group
- Numerical Columns: Age, Minutes Streamed Per Day, Number of Songs Liked, Discover Weekly Engagement (%), Repeat Song Rate (%)

EDA

Univariate Analysis

- Histograms and Boxplots for numerical columns
- Count Plots for categorical columns
- Mean, Median, Mode, Standard Deviation, etc. for numerical columns
- Mode, Frequency, etc. for categorical columns
- Distribution of data in each column

Bivariate Analysis

- Correlation matrix to identify relationships between numerical variables
- Scatter plots for continuous variable relationships
- Bar plots, violin plots, and box plots to compare categorical and numerical variables

Multivariate Analysis

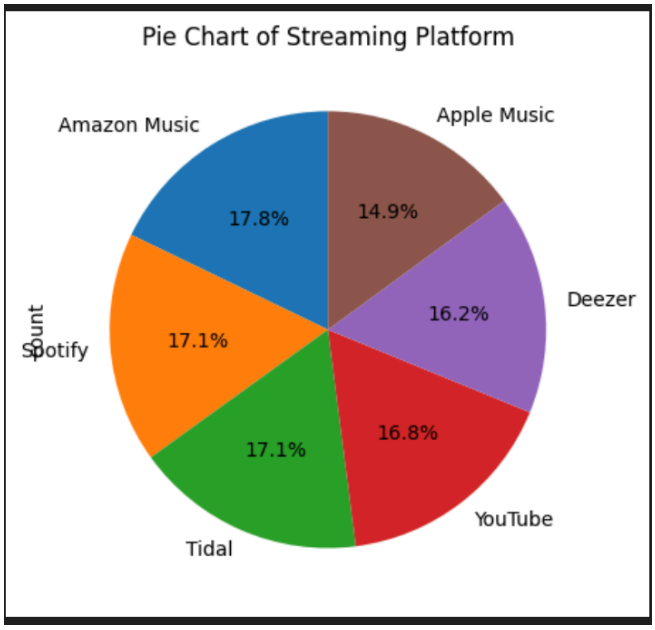
- Pair plots to analyze multiple relationships simultaneously
- Heatmaps to visualize correlations among multiple variables

Group Comparisons - Very Important for Music Industry Companies

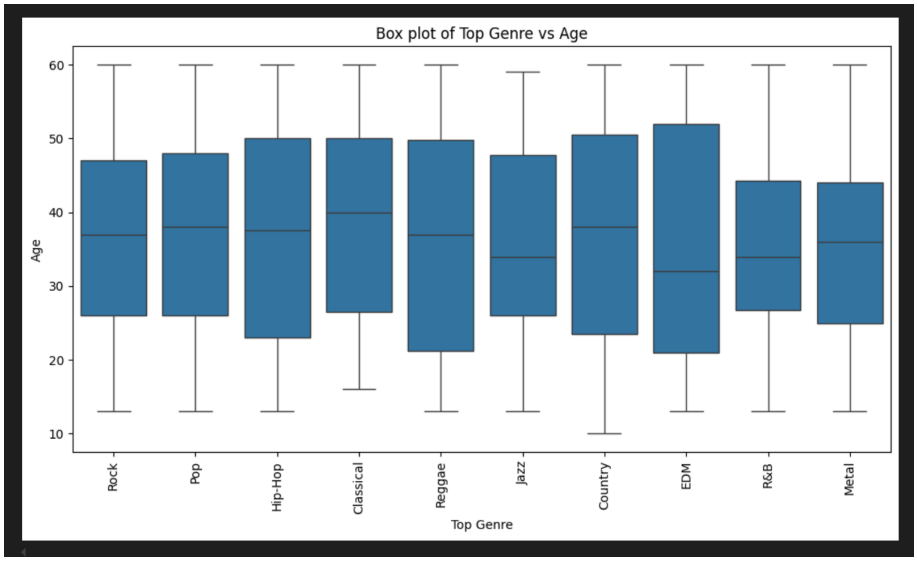
- Comparing listening time, platform, subscription, etc based on country and age group
- Understanding the preferences of users based on their country and age group

Key Conclusions from the Music Streaming Dataset :

1. **Streaming Platform Choices**
    - All platforms are equal in terms of number of users, so can't say anyone is more popular.



2. **Age Group & Music Preferences**
    - Young adults (19-25) stream the most music daily, averaging 400+ minutes per day.
    - Older users (46+) prefer Classical and Jazz, while Hip-Hop and Pop dominate younger audiences.



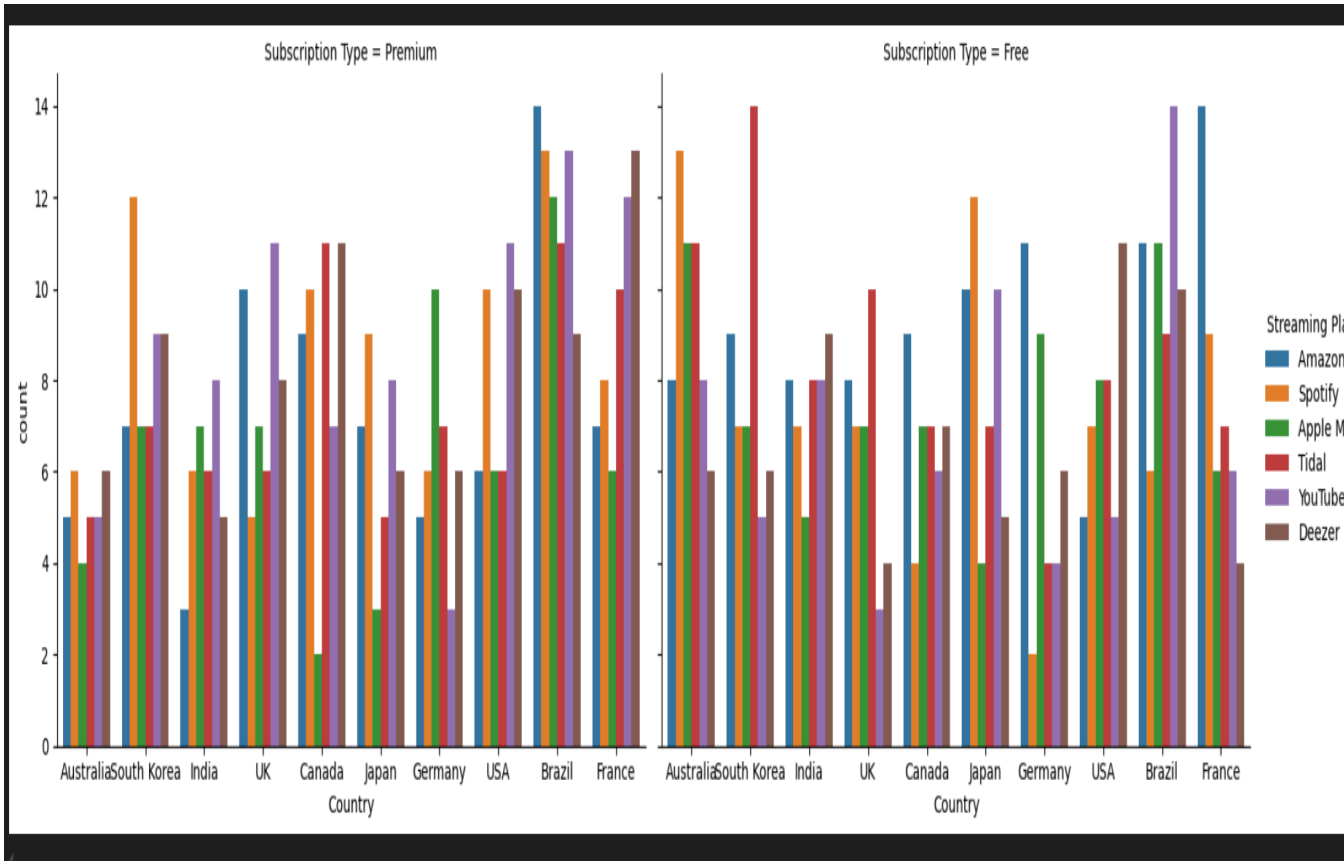
3. **Countries vs Platform and Subscription type**

Bar chart is used to compare the distribution of streaming services among various countries for Premium and Free users.

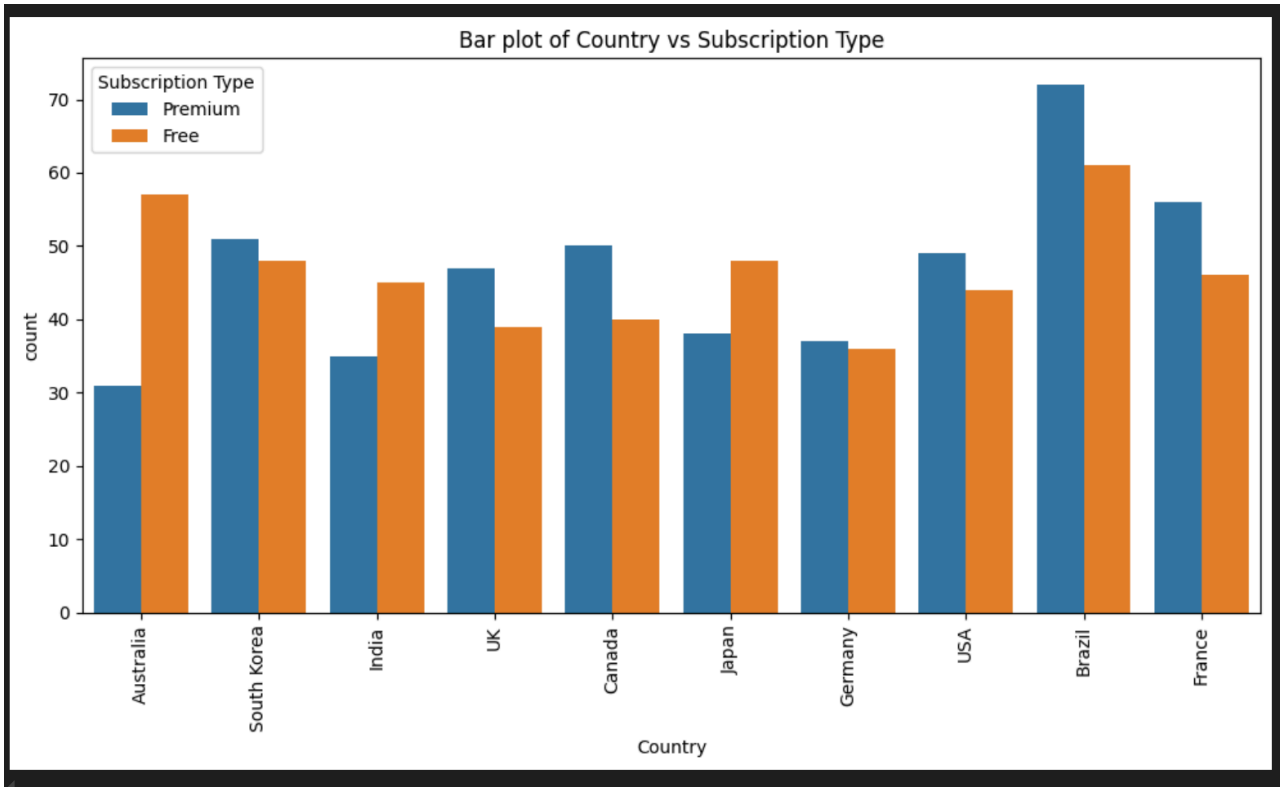
Premium Users:

Most countries are dominated by Spotify, particularly South Korea and Japan.
YouTube and Deezer boast a large Premium base in Brazil.
Amazon Music is well-known in the UK and USA and is a strong competitor of Spotify.
Free Users:

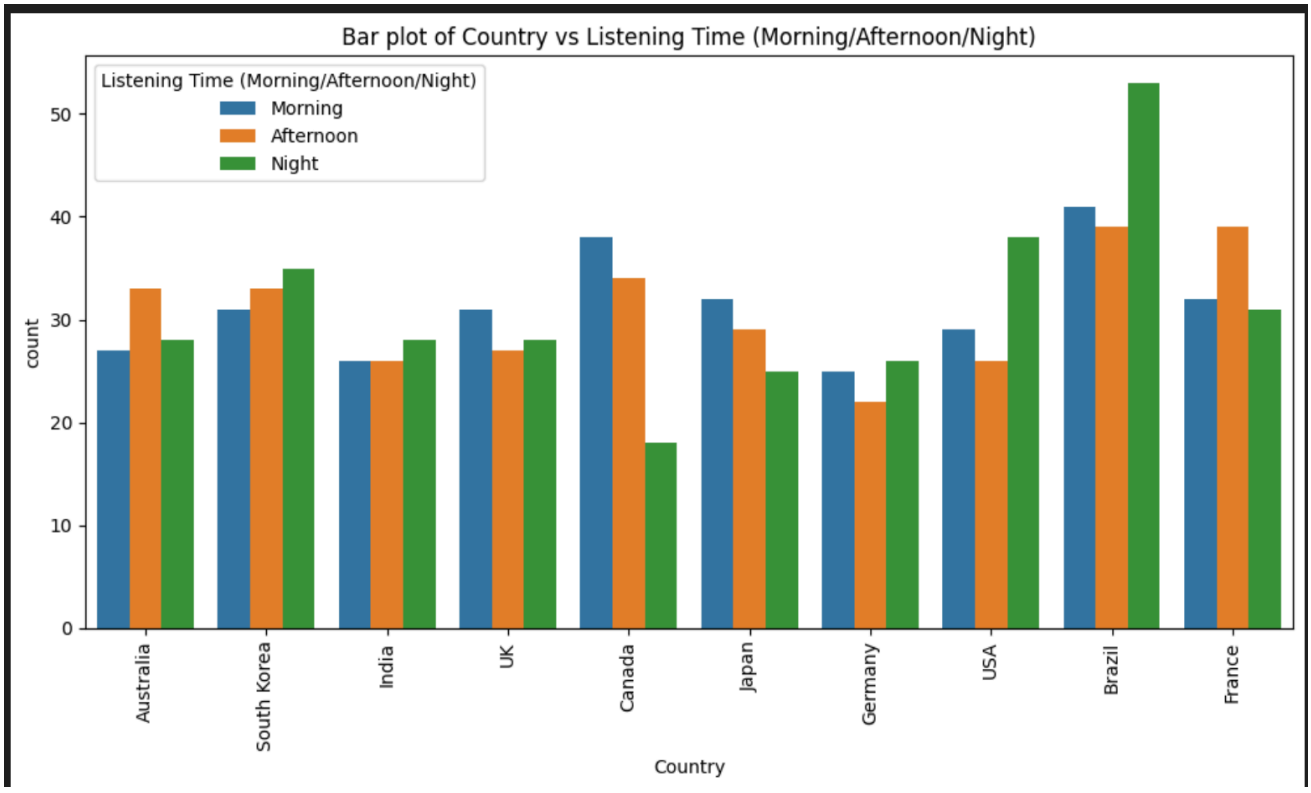
YouTube has become more popular among free users, particularly in Brazil and France.
Tidal has an exceptionally high number of free users in South Korea, implying local affinity.
Spotify continues to be robust but experiences increased competition from YouTube and Deezer in a few countries.
Insight: Spotify dominates among premium users, whereas YouTube and Deezer draw more free users. Country-specific preferences are diverse, with Brazil, South Korea, and India depicting distinct streaming patterns.



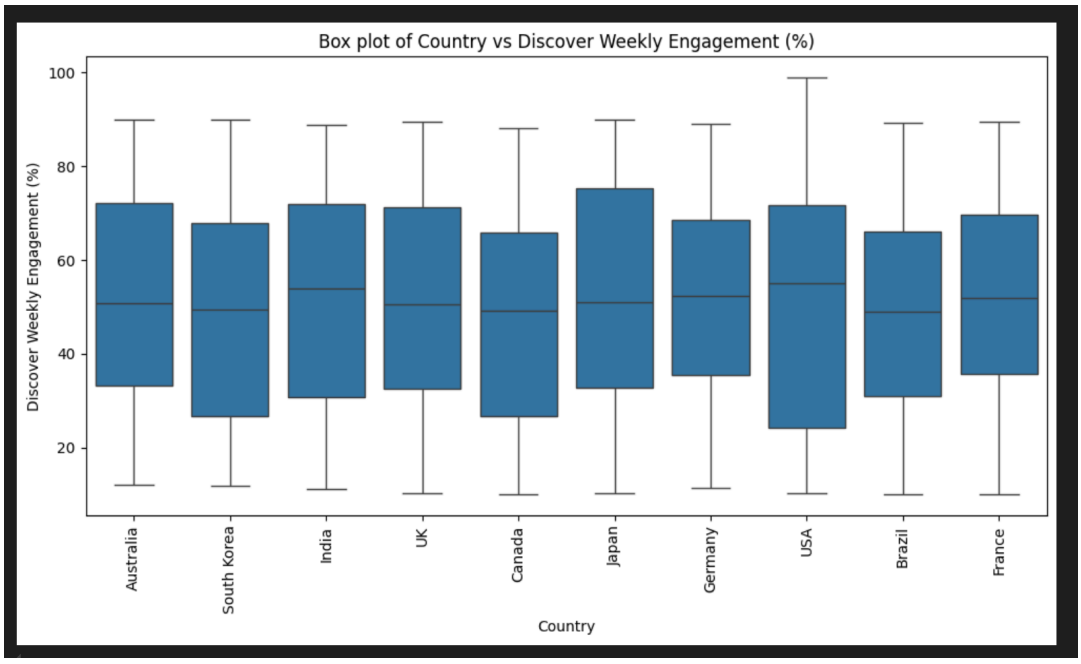
4. **Country vs Subscription**
    - By analyzing it companies can focus on the country's which prefer more subscriptions,
    - Brazil and the USA have the most premium users, while Australia and Japan prefer free subscriptions, according to the bar plot comparison of subscription types across nations. The distribution of premium and free users is balanced in the majority of other nations.



5. **The distribution of listening times** (morning, afternoon, and night) in various nations is depicted by the bar plot. Canada exhibits a strong preference for morning listening, whereas Brazil has the highest nighttime listening rates. With a few exceptions where afternoon listening is marginally more popular, the distribution is generally balanced in the majority of other nations.



6. The box plot displays the distribution, median, and variability of **Discover Weekly engagement** percentages in various nations. The range of engagement levels in the majority of nations is comparable, ranging from low to high. While other nations show similar trends with a few outliers, Japan and the USA seem to have somewhat higher medians.



Concluding it !

- Numerous insights regarding user engagement in various nations are revealed by the data.
- The distribution of subscription types reveals that the number of Premium and Free users varies by nation, with the highest Premium users found in Brazil and the USA.
- According to listening time patterns, preferences vary by time of day, with some nations favoring listening in the morning or at night.
- Find out Weekly engagement shows a wide range of participation, with more balanced but varied distributions in other countries and comparatively higher medians in the USA and Japan.
- And lots and lots of takeaways we can get from the dataset !

Final Business Takeaway:

- Because they have the highest engagement, companies should concentrate on the USA and Brazil for Premium subscriptions.
- Discover Weekly engagement is higher in countries like the USA and Japan, which makes them perfect for exclusive content strategies and tailored recommendations. Due to the fact that listening
- Different countries have different listening habits, so marketing strategies should be adjusted accordingly. For example, some countries have strong night listeners, while others prefer mornings.
- Discover Weekly's lower engagement areas might benefit from incentives or promotions to increase interest, whereas areas with a high rate of free subscriptions might need different monetization strategies, like advertisements or freemium benefits.
- Hence concluding everything, there are many conclusions that companies can conclude and make strategies to improve their position in the Music industry.