Project Title: **CASE STUDY 2**

Analysis of Spotify Data

Arnav Jain

OMT 548: Statistical Methods for Rsrch

Prof. Jason Bronowitz

5th May, 2023

Story

Once upon a time, there was a music lover named Sarah who was a data enthusiast as well. Sarah used Spotify to listen to music while studying, working out, and commuting. One day, she decided to analyze her streaming history data to understand her music listening habits.

Sarah began her analysis by visualizing her total hours of music playback per week. She discovered that her music listening activity varied widely, with some weeks having very little music playback, while others had much more. Sarah realized that her streaming activity was influenced by factors like her workload, social events, and overall mood.

Next, Sarah examined her top 10 most streamed artists and songs. She found that she listened to a diverse range of genres, with pop, hip-hop, and rock being her most frequently played genres. Sarah was surprised to find that some of her favorite songs were not by her top-listed artists, suggesting that she was exploring new music as well.

Further analyzing her data, Sarah examined her playback activity by time of the day. She discovered that she tended to listen to music during the afternoons and evenings on weekdays, while weekends were more evenly distributed throughout the day. She realized that her music listening habits were a reflection of her daily routine and schedule.

Through her analysis, Sarah had gained valuable insights into her music listening habits. She realized that her music tastes were diverse, her listening habits were influenced by her daily routine, social life, and mood, and that she was always looking for new music to enjoy. Sarah felt a greater appreciation for the music she listened to and continued to explore new artists and genres on Spotify.

PROJECT WRITEUP

Data validation-

The data contains a Json file. This report presents the data validation results of Spotify data containing 6866 entries with four columns: Artist name, trackname, msPlayed, and endTime. The validation aimed to ensure that the data is accurate, complete, and consistent with the expected format.

Data Includes-
 Endtime- It is the time and date data starting from 01-01-20 to 01-17-21

Artist Name- Contains data of each artist that Sarah listened to between 01-01-20 to 01-17-21. It contains 423 different artists.

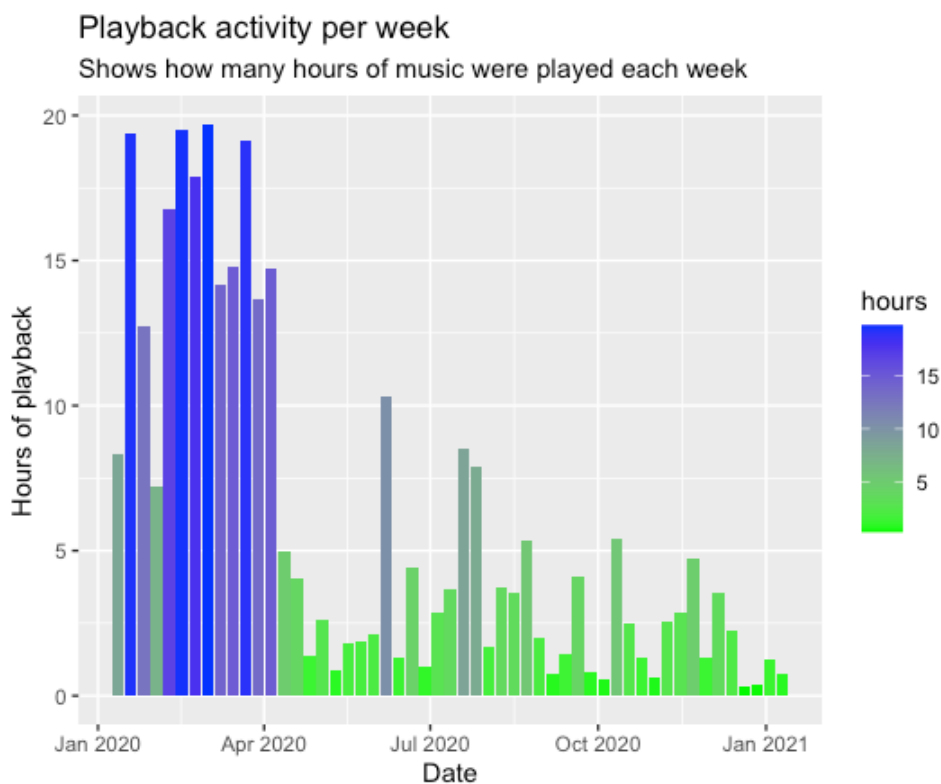Track name- Contains data of each track listened by Sarah. It has 6866 Tracks
msPlayed- Contains data of Milli seconds played per track.
After going through the data we checked –

- Checking for missing values: We checked if there were any missing values in any of the columns. We found that there were no missing values in the data.

- Checking for duplicates: We checked if there were any duplicate entries in the data. We found that there were no duplicate entries in the data.

- Checking for invalid data: We checked if the data was in the expected format. We found that all the columns were in the expected format. The artist name and song columns only contained alphabetical characters and spaces, and the msPlayed column only contained numerical values.
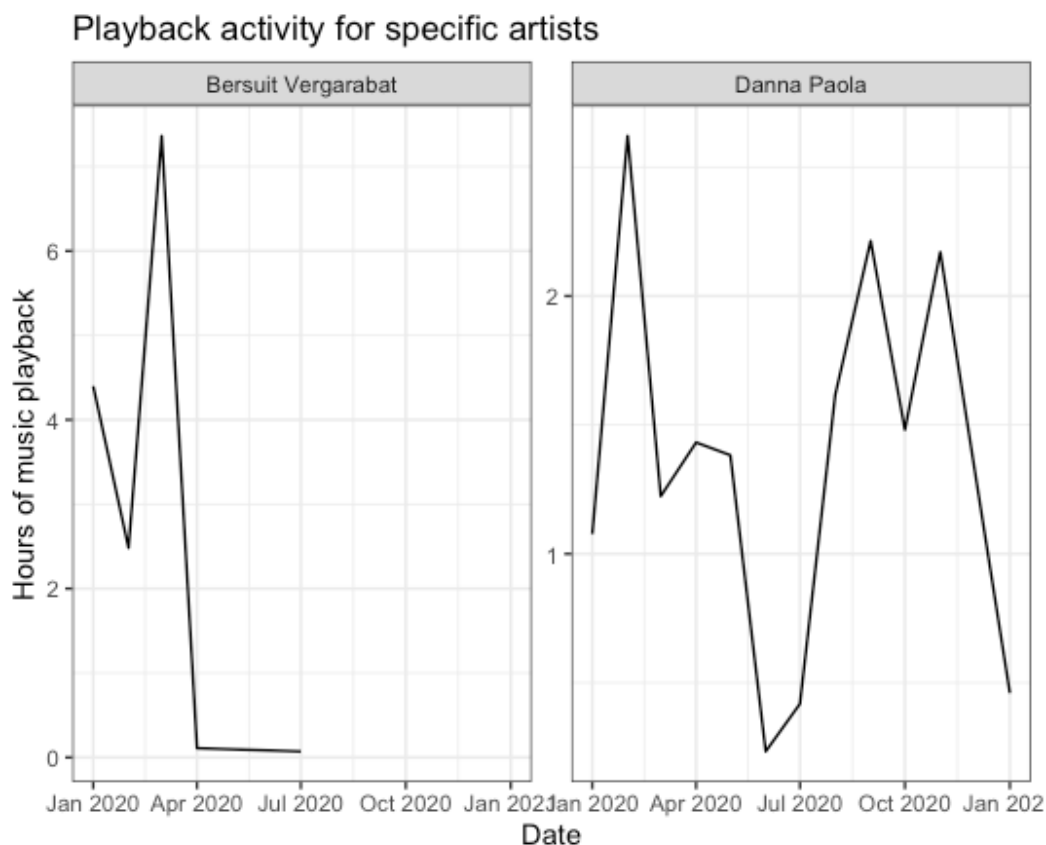
Exploratory Analysis-

1. Playback Activity per week-

Playback activity per week
Shows how many hours of music were played each week

Through our data visualization, Upon examining Sarah's music playback activity on a monthly basis, it was observed that she listened to a significant amount of music in January and April, with peak listening periods occurring in March and February. This pattern indicates that Sarah's music consumption was higher during the early months of the year and gradually declined in the later part of the year.

2) Analysis of Playback of top 2 artists-

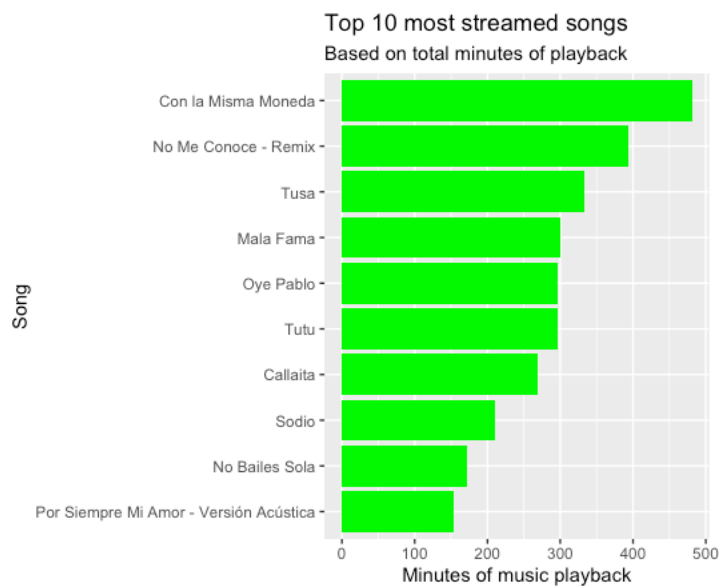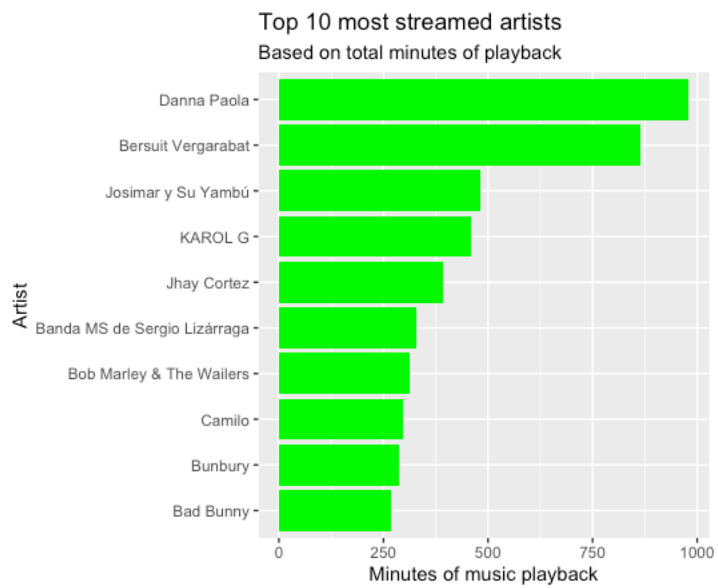### Playback activity for specific artists

In the exploration of Sarah's music streaming history, a fascinating pattern emerged when examining her playback activity for her top two artists: Bersuit and Danna Paola.

For the artist Bersuit, Sarah's listening behavior showcased a distinctive trend. From January to April, Sarah's engagement with Bersuit's music experienced a noticeable increase, reaching its peak in February. This suggests a heightened interest and connection with Bersuit's music during this period. It is worth exploring the possible reasons behind this surge in Sarah's listening activity. Perhaps Bersuit released new music, or their songs resonated with Sarah during that particular time frame.
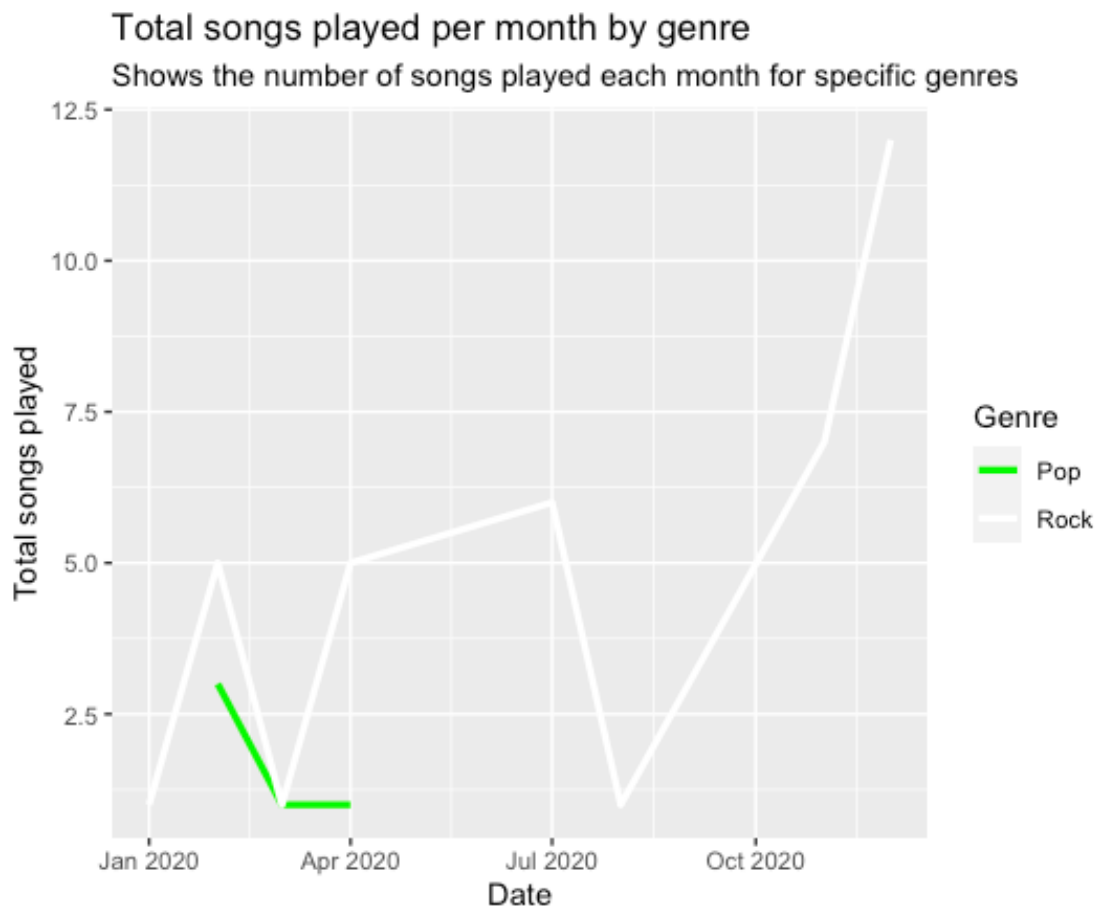
On the other hand, with Danna Paola, Sarah consistently listened to her music throughout the entire year. This finding indicates a strong and enduring affinity for Danna Paola's musical style and compositions. Danna Paola's music likely holds a special place in Sarah's heart, as she regularly turned to her songs for enjoyment, inspiration, or emotional connection

3) Top 10 Artists and Songs-

Top 10 most streamed artists
Based on total minutes of playback

Top 10 most streamed songs
Based on total minutes of playback

Here includes the list of Sarah's top 10 Songs and artists. This suggests that Sarah is mostly into Latin Music.

4) Top Genres-

Total songs played per month by genre
Shows the number of songs played each month for specific genres

Sarah's love for pop music was strong in the beginning. She listened to popular pop songs and discovered new artists in that genre. However, as time went on, especially in April, she seemed to lose some interest in pop music. It could be a temporary phase or a change in her musical preferences during that time.
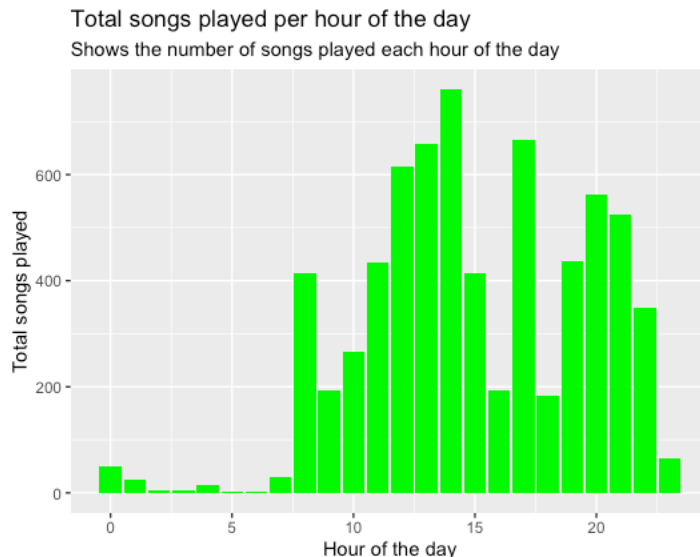
On the other hand, Sarah's fondness for rock music grew steadily throughout the year. While her interest in pop declined, she found herself exploring and appreciating rock music more and more. Towards the end of the year, her excitement for rock music reached its peak. This indicates that Sarah began to delve into the world of rock, discovering new rock artists, albums, and songs that resonated with her.

Sarah's evolving taste in music highlights her openness to different genres. It's possible that she wanted to explore and expand her musical horizons. This information is valuable for music platforms and artists as they can curate personalized recommendations and playlists that cater to Sarah's changing interests. By offering a wide variety of rock music content, they can continue to engage and captivate Sarah's attention throughout the year.
In summary, Sarah started the year with a strong interest in pop music but experienced a decline in April. However, her passion for rock music steadily grew and peaked by the end of the year. This shift in genre preference demonstrates Sarah's willingness to explore different

musical styles and provides insights for music platforms and artists to create personalized experiences for her

4) Total Songs played per hour-

Sarah's music listening habits reveal a clear pattern throughout the day. She showed a higher level of engagement with music during the afternoon hours, particularly peaking at 2 pm. This indicates that Sarah enjoyed indulging in her favorite songs and artists during that time, possibly during her break or leisure hours.

As the day progressed towards the evening, Sarah's music consumption gradually decreased. After 8 pm, there was a noticeable drop in her music listening activity. This suggests that she might have shifted her focus to other activities or simply preferred a more relaxed evening without heavy music exposure.

Recommendations-

Based on the data and analysis presented in the story, here are some recommendations:
1. Keep exploring new music: Sarah discovered that she was listening to a diverse range of genres and was always looking for new music to enjoy. This is a great habit to maintain, as it helps to broaden one's musical horizons and discover new artists and sounds.
2. Pay attention to mood and routine: Sarah realized that her music listening habits were influenced by her daily routine, social life, and mood. It may be helpful to pay attention to these factors when selecting music to listen to, as different types of music can have a positive impact on mood and energy levels.
3. Share music with friends: Sarah found that she had discovered new artists through her friends and social media. Sharing music with friends can be a great way to discover new artists and expand one's musical taste.

References-

1) https://github.com/cosmoduende/r-spotify-history-analysis/blob/main/StreamingHistory0.json
Got the data from Here.