GROUP 3 SPOTIFY PROJECT PROPOSAL

Bri Abreu and Maria Kraft

https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-01-21/readme.md

BACKGROUND.

As passionate German techno lovers aiming to establish a production company in the vibrant Atlanta music scene, we recognize the importance of understanding local music preferences and trends. Our objective is to leverage Spotify data to gain insights into the Atlanta audience's musical inclinations, enabling us to tailor our techno productions to resonate with the local taste.

MOTIVATION.

PROBLEM STATEMENT

The challenge is to bridge the cultural gap between German techno, known for its unique characteristics, and the diverse music scene in Atlanta. We need to identify specific elements within our techno genre that align with the preferences of the Atlanta audience, ensuring our productions gain popularity and acceptance in this dynamic market.

RESEARCH QUESTION

"How can we adapt our German techno productions to align with the musical preferences of the Atlanta audience, ensuring a successful integration into the local music scene?"

WE STARTED WITH 32,833 ROWS OF DATA AND WERE LEFT WITH 23,449 AFTER CLEANING.

DATASET CHARACTERISTICS.

There are 4981 unique track names that have duplicates, resulting in a total of 14365 rows with duplicate entries.

Dependent Variable:

Track Popularity: This variable, ranging from 0 to 100, will serve as our primary dependent variable. We hypothesize that the popularity of a track is influenced by various independent variables.

Independent Variables:

Danceability: We hypothesize that tracks with higher danceability scores will be associated with higher popularity, assuming that danceable tracks are more likely to be well-received by a diverse audience.

INDEPENDENT VARIABLES CONT'D.

Energy: Higher energy levels are expected to correlate positively with track popularity. Energetic tracks often capture attention and engage listeners, potentially contributing to higher popularity scores.

Speechiness: We anticipate that tracks with moderate speechiness (indicating a blend of music and spoken words) may have higher popularity, catering to a wider audience that enjoys a mix of musical and lyrical elements.

Acousticness: While techno typically has lower acousticness, we expect that tracks with some acoustic elements might resonate well in certain playlists, potentially influencing popularity.

Instrumentalness: Higher instrumentalness is expected to correlate positively with track popularity, as techno, being an instrumental genre, might attract listeners who appreciate tracks without prominent vocals.

Valence: Positive valence is likely to contribute to higher popularity. We hypothesize that tracks conveying positive emotions align with the upbeat and lively nature of the Atlanta music scene.

DATASET CHARACTERISTICS CONT'D .

Liveness: We expect that a higher liveness value, indicating a live performance, may positively influence popularity, adding an authentic and dynamic element to the track.

Tempo: The overall estimated tempo may impact popularity, as the Atlanta music scene may favor certain BPM ranges. We anticipate a correlation between moderate tempos and higher popularity.

Hypothesis

By systematically exploring these relationships and considering the contextual relevance of the variables in the Atlanta music scene, we aim to draw meaningful conclusions to guide our production strategies effectively. We hypothesize that pop and rap elements will allow us to reach more listeners.

APPROACH & METHODOLOGY.

GENRES

POP

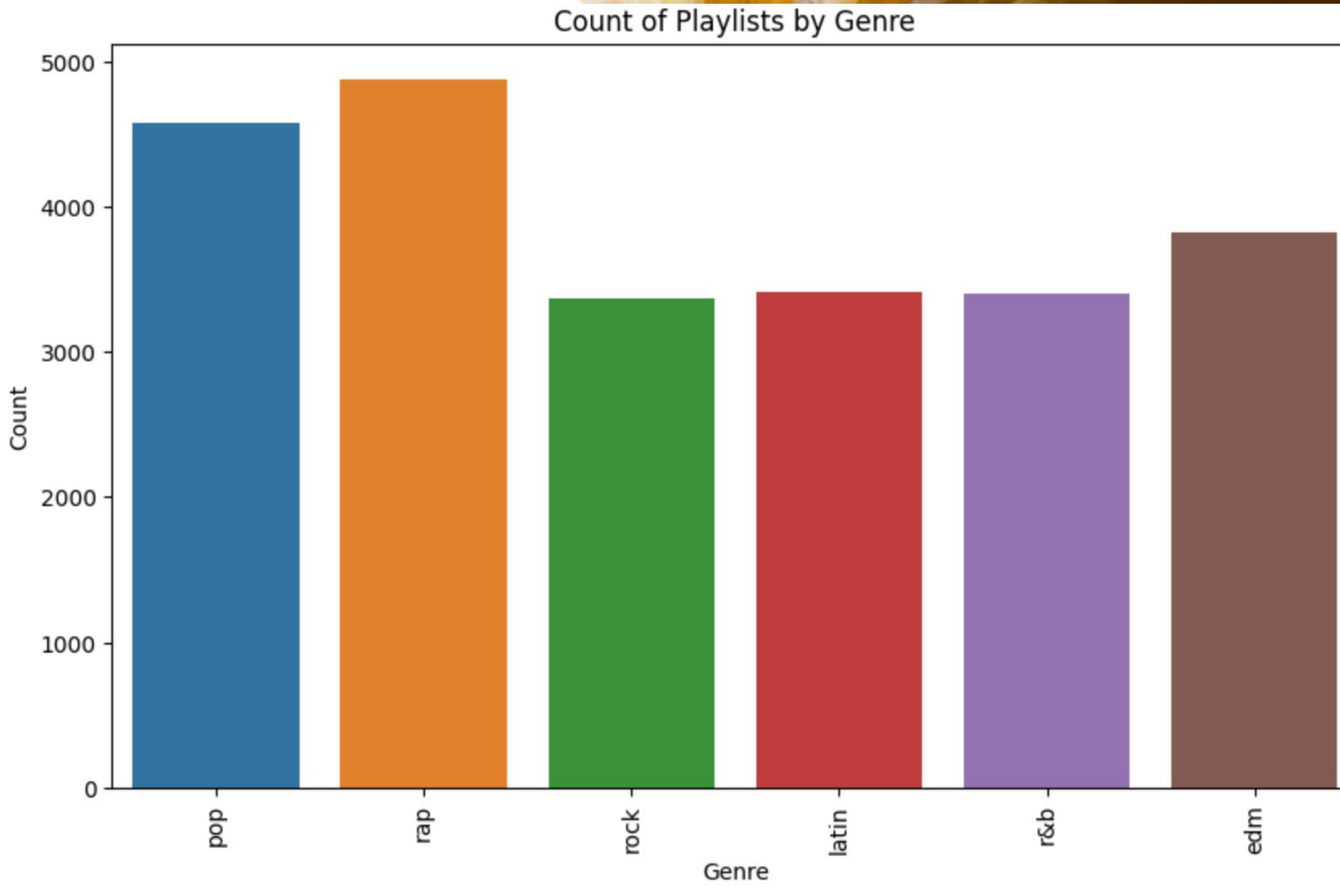
RAP

ROCK

LATIN

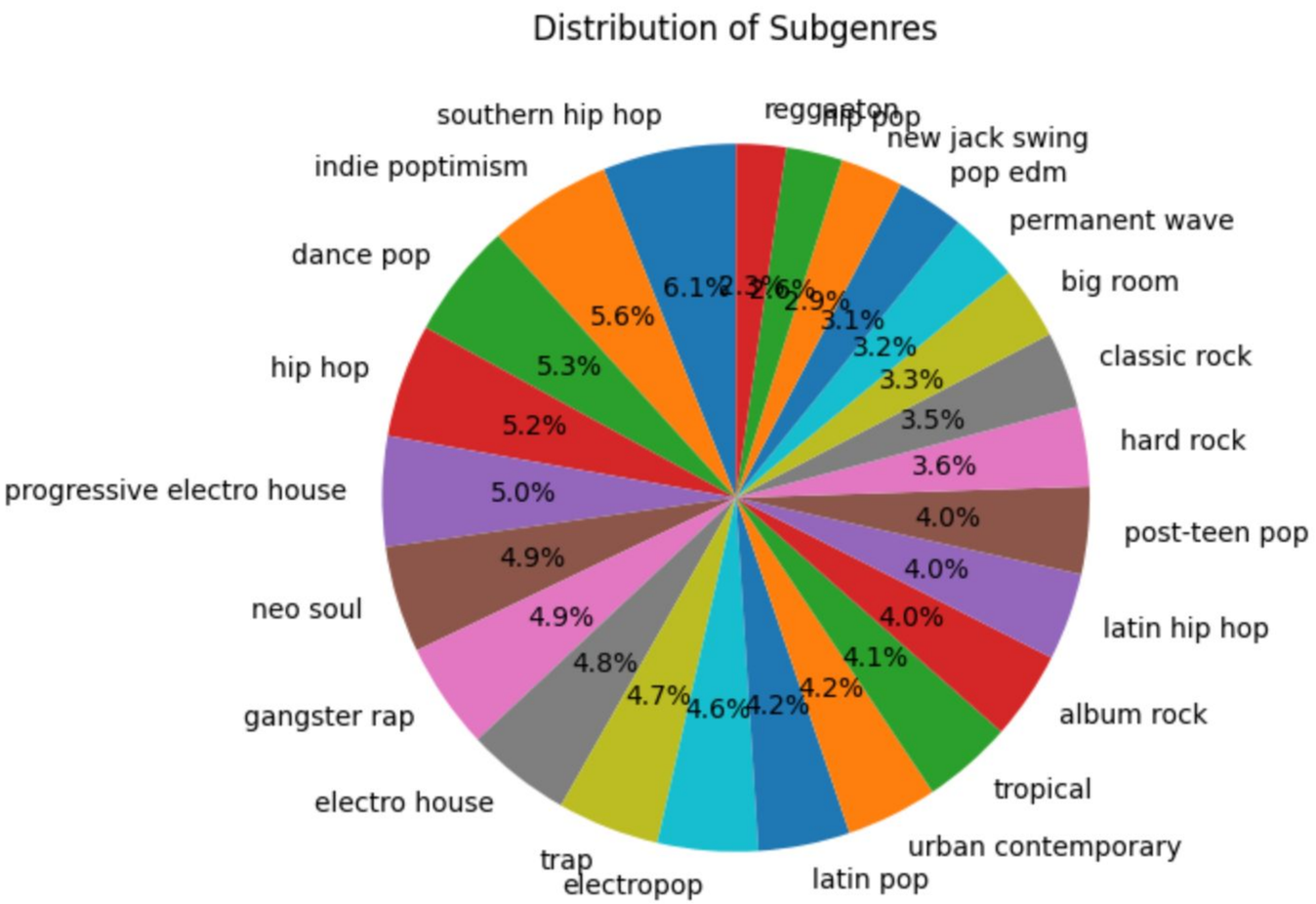
R&B

EDM

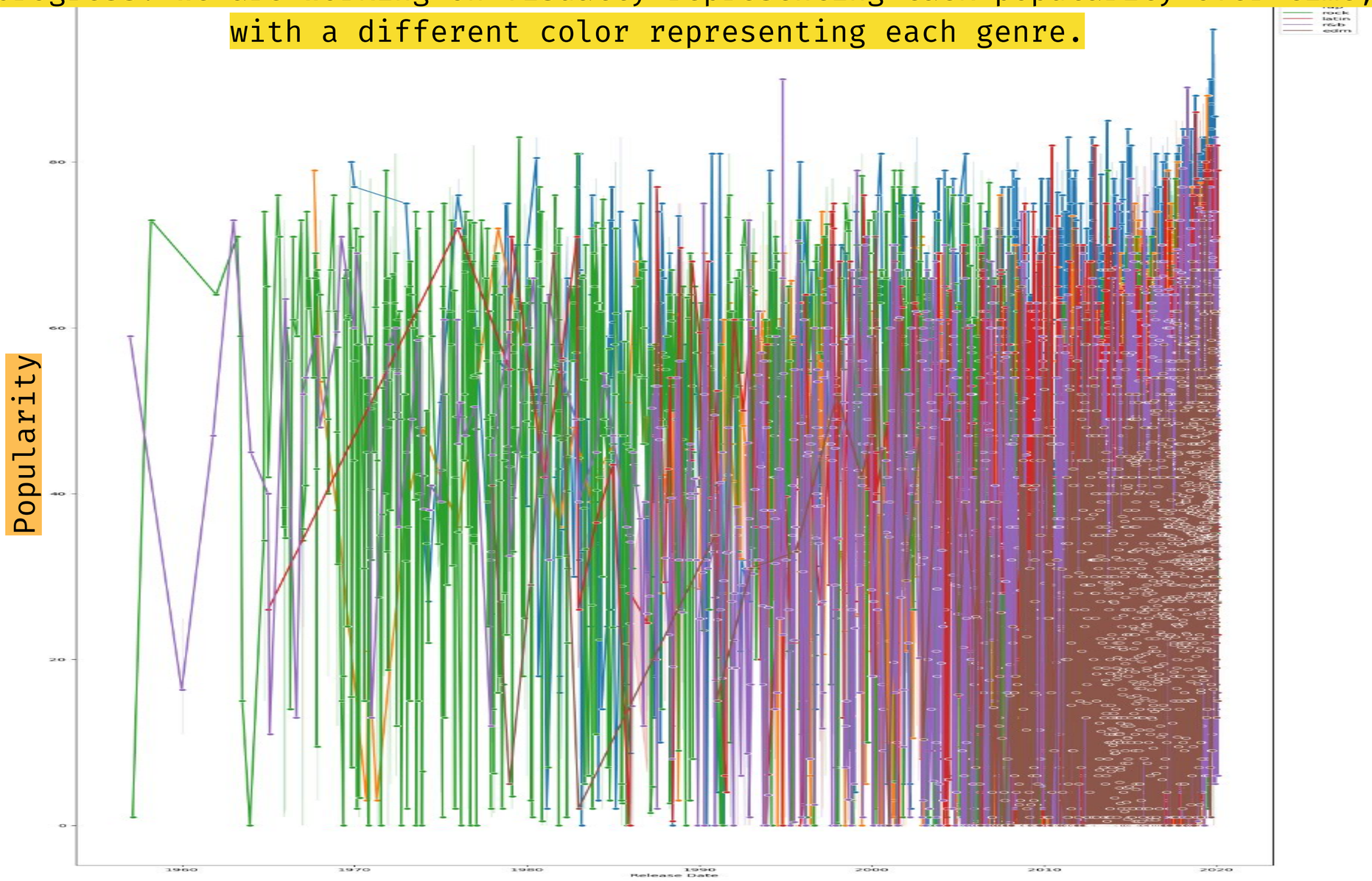


- Rap is the most popular genre

- Southern Hip Hop is the most popular subgenre



In progress: We are working on visually representing tack popularity over time, with a different color representing each genre.



Release Date

PROPOSED APPROACH .

- We will conduct a correlation analysis between the dependent variable (Track Popularity) and each independent variable to identify initial associations.
  - Regression models to understand the strength and direction of relationships between the dependent and independent variables, controlling for potential confounding factors.

- Visualizations: Create visualizations such as scatter plots, bar charts, and line graphs to illustrate the relationships between key variables and track popularity.

Using our findings, we will determine the “perfect” song characteristics. We will then produce our EDM tracks to match the taste of the listeners.

For example, we found that rap is the most popular genre. We can incorporate clips of popular rap songs in our tracks to increase plays.

Next Steps:

- Additional data exploration
- Creation of new variables if needed

WRAP UP

Production Strategy: The analysis will guide our techno production strategy, ensuring that our music aligns with the musical preferences of the Atlanta audience. This will enhance the likelihood of local acceptance and popularity.

Marketing Tactics: By understanding the popular genres, subgenres, and playlist dynamics in Atlanta, we can optimize our marketing tactics. This includes targeted advertising, collaborations, and strategic playlist placements to maximize visibility and engagement.