Data Science Project Report: Analyzing Seasonal Music Preferences

1. Introduction

1.1 Motivation

The motivation behind this project is to explore and understand how music preferences evolve across different seasons. We aim to investigate whether there are noticeable patterns in the choice of genres and audio features, and how they may be influenced by seasonal changes.

2. Data Collection and Cleaning

2.1 Data Source

The primary data source for this project is a personal Spotify listening dataset. The dataset includes information on music tracks, such as artist name, track name, end time, and the ms played each song listened on Spotify.

2.2 Data Collection

The data was collected via requesting Spotify to obtain the personal data. Furthermore, data was enriched by leveraging the Spotify API to retrieve detailed audio features, and genres for each track. The dataset was then cleaned using the Pandas library in Python, addressing issues like missing values and formatting inconsistencies.

3. Data Analysis

3.1 Hypotheses

Null Hypothesis (H_o):

"There is no significant difference in music preferences, as measured by audio features and genres, across different seasons. Specifically, the mean values of both audio features and genres remain constant throughout each season."

Alternative Hypothesis (H_a):

"There is a significant difference in music preferences, as measured by audio features and genres, across different seasons. Specifically, the mean values of

either audio features or genres (or both) change noticeably from one season to another, indicating a seasonal variation in music preferences."

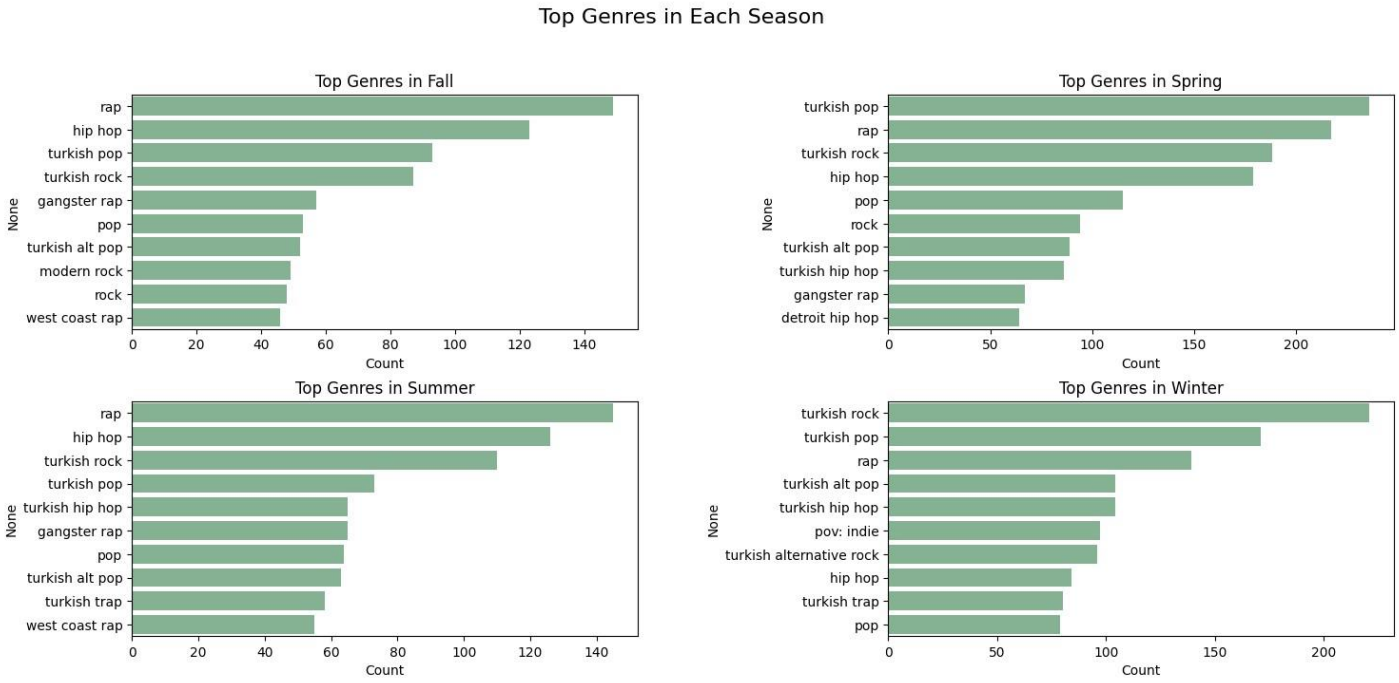
3.2 Techniques Used

- Exploratory Data Analysis (EDA)
- Feature Extracting via API calls
- Data Visualization using Seaborn, Matplotlib
- Statistical analysis to test hypotheses

4. Data Visualizations

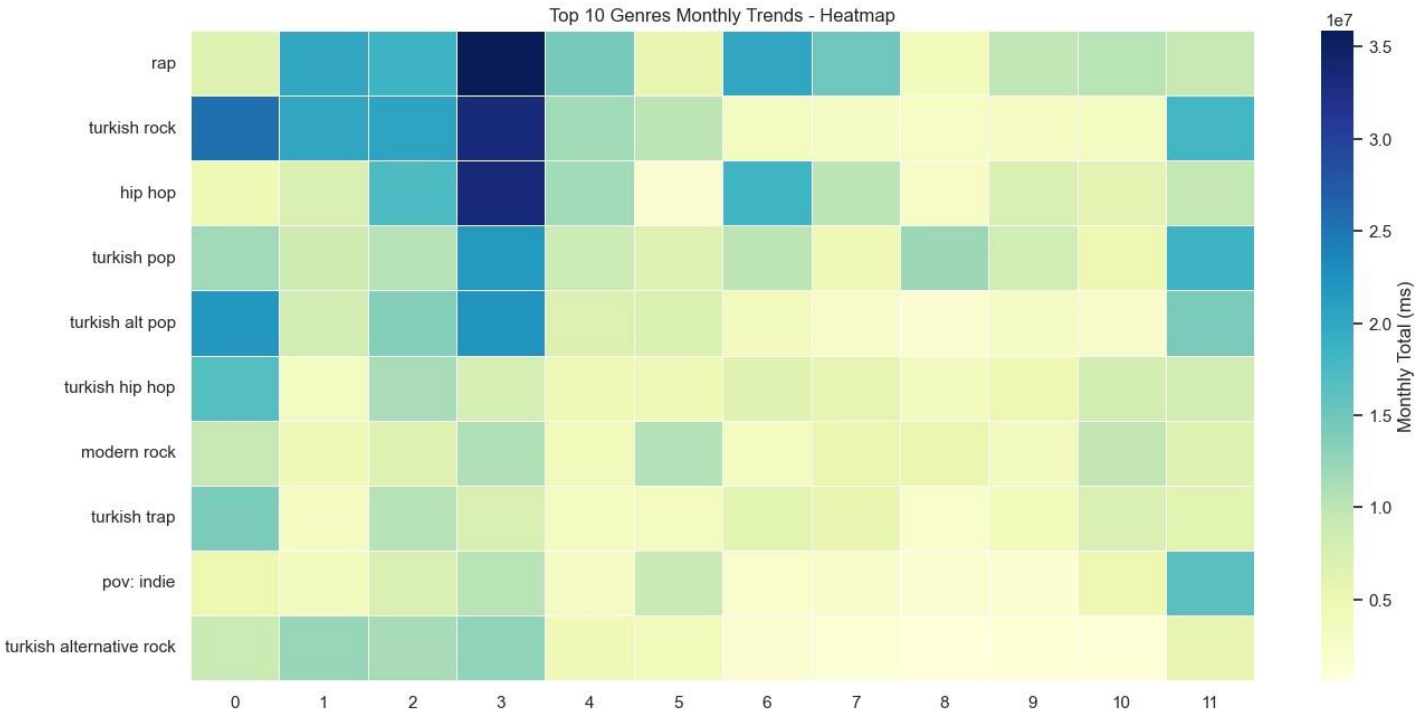
4.1 Top Genres in Each Season

We present bar charts showcasing the top genres listened to in each season, providing insights into seasonal genre preferences.



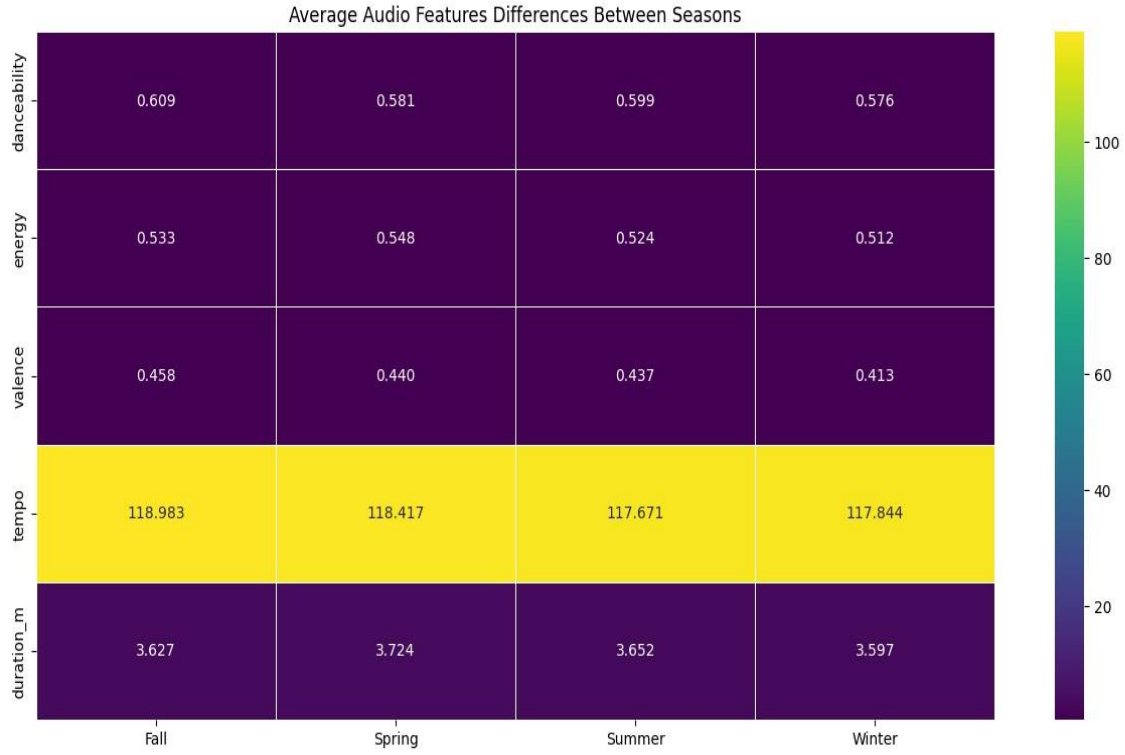
4.2 Top 10 Genres Monthly Trends - Heatmap

A heatmap illustrates the monthly trends of the top 10 genres, allowing us to identify patterns and shifts in genre preferences over time.



4.3 Average Audio Features Differences Between Seasons

A heatmap depicts the average differences in audio features (e.g., danceability, energy) between seasons, highlighting variations in musical characteristics across the year.



5. Findings

Our analysis reveals notable patterns in music preferences, with certain genres dominating specific seasons. However, the heatmap analysis does not demonstrate significant differences in average audio features between seasons.

6. Limitations and Future Work

6.1 Limitations

- The dataset is based on personal listening habits and may not be representative of general trends.
- The Spotify API provides a snapshot of data, and preferences may evolve over time.

6.2 Future Work

- Incorporate larger and more diverse datasets for a broader understanding of seasonal music preferences.
- Explore external factors (e.g., weather, holidays) to identify additional influences on music choices.

7. Conclusion

"The results of the data analysis suggest that there is a significant difference in music preferences in terms of genres across different seasons. This finding supports the alternative hypothesis (H_a) that there is a noticeable variation in the mean values of genres. However, it is observed that the mean values of audio features remain constant throughout each season, which does not provide enough evidence to reject the null hypothesis (H_o) regarding audio features.

"In summary, while there is a observed seasonal variation in music preferences, particularly in genres, the audio features exhibit a notable consistency across different seasons. Also these audio features reveals a preference for music that is neither overwhelmingly energetic nor subdued, and with a danceability level indicating a moderate inclination towards tracks suitable for dancing."