Analysis of Spotify's Top 100 Songs from 2010 to 2019

1. Introduction:

Spotify, a premier music streaming platform, witnessed numerous tracks gaining immense popularity over the past decade. This analysis delves into the top 100 songs from 2010 to 2019, examining attributes such as song energy, danceability, and popularity. The dataset offers a glimpse into the musical predilections of Spotify's listeners during this era, aiding in the identification of music industry trends.

Dataset Overview:

The dataset encompasses the following columns:

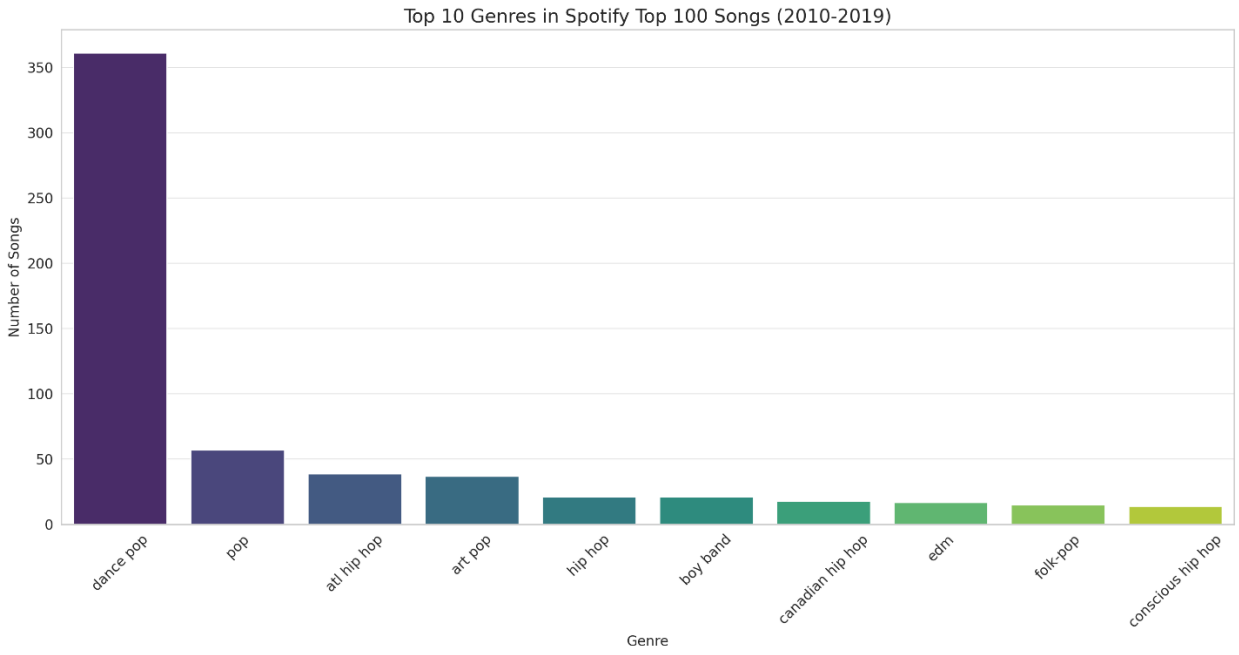
1. **Title:** Song title.
2. **Artist:** The artist or group that performed the song.
3. **Top Genre:** The primary genre of the song.
4. **Year Released:** The song's release year.
5. **Added:** The date when the song was included in the dataset.
6. **BPM:** Beats per minute, indicative of the song's tempo.
7. **Energy (nrgy):** A metric of intensity and activity.
8. **Danceability (dnce):** Illustrates a track's suitability for dancing.
9. **dB:** The song's loudness.
10. **Live:** Indicates if a live audience is present in the recording.
11. **Valence (val):** A metric showcasing the musical positivity of a track.
12. **Duration (dur):** Song length in seconds.
13. **Acousticness (acous):** A measure indicating the likelihood of the track being acoustic.
14. **Speechiness (spch):** Detects spoken words in a track.
15. **Popularity (pop):** The song's popularity score.
16. **Top Year:** The year the song ranked in the top 100.
17. **Artist Type:** Specifies if the artist is a solo performer or part of a duo/group.

Visualizations and Summaries:

1. Distribution of Songs by Genre

Code:

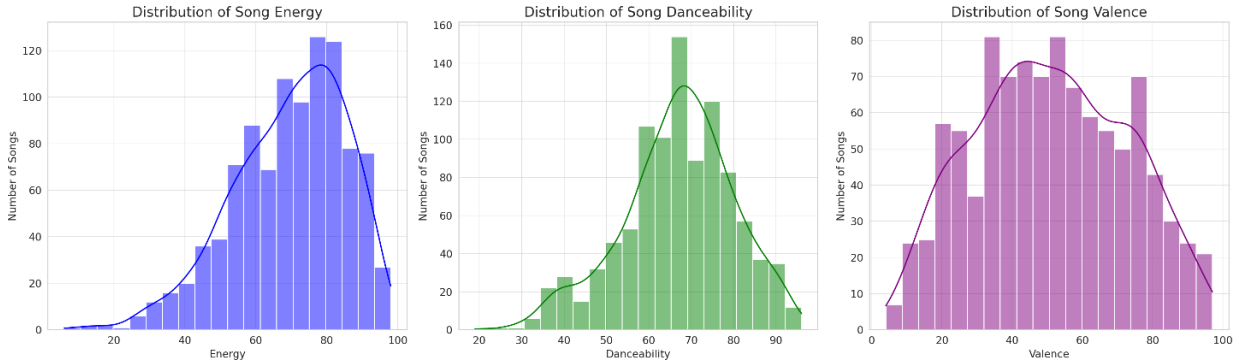
```
spotify_data_csv = pd.read_csv("/mnt/data/Spotify 2010 - 2019 Top 100.csv")
sns.set_style("whitegrid")
plt.figure(figsize=(15, 8))
genre_counts = spotify_data_csv['top genre'].value_counts().head(10)
sns.barplot(x=genre_counts.index, y=genre_counts.values, palette='viridis')
plt.title('Top 10 Genres in Spotify Top 100 Songs (2010-2019)')
plt.xlabel('Genre')
plt.ylabel('Number of Songs')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



Summary:
The bar plot displays the distribution of the leading 10 genres in Spotify's Top 100 songs between 2010 and 2019. Dance Pop emerges as the dominant genre, emphasizing its widespread acceptance throughout the decade. Pop also features prominently, but Dance Pop's dominance is clear.

2. Distribution of Song Attributes - Energy, Danceability, Valence
Code:

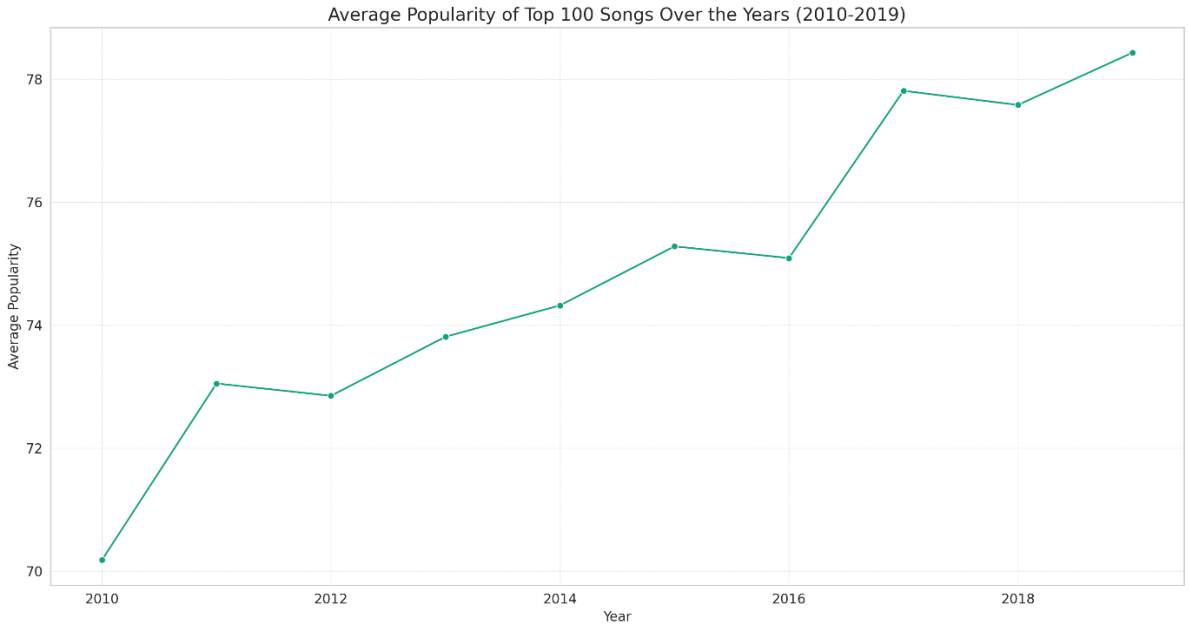
```
f, axes = plt.subplots(1, 3, figsize=(20, 6))
sns.histplot(spotify_data_csv['nrgy'], kde=True, ax=axes[0], color='blue', bins=20)
axes[0].set_title('Distribution of Song Energy')
axes[0].set_xlabel('Energy')
axes[0].set_ylabel('Number of Songs')
sns.histplot(spotify_data_csv['dnce'], kde=True, ax=axes[1], color='green', bins=20)
axes[1].set_title('Distribution of Song Danceability')
axes[1].set_xlabel('Danceability')
axes[1].set_ylabel('Number of Songs')
sns.histplot(spotify_data_csv['val'], kde=True, ax=axes[2], color='purple', bins=20)
axes[2].set_title('Distribution of Song Valence')
axes[2].set_xlabel('Valence')
axes[2].set_ylabel('Number of Songs')
plt.tight_layout()
plt.show()
```



Summary:
These histograms depict the distribution of key song attributes: Energy, Danceability, and Valence. The majority of top-ranked songs possess high energy levels, peaking around the 80 mark. Danceability presents a broader distribution, but a discernible peak is observed around 70. Valence, representing a song's positivity, demonstrates a wide-ranging distribution, signaling the diverse emotional inclinations of listeners.

3. Average Popularity of Songs Over the Years
Code:

```
plt.figure(figsize=(15, 8))
sns.lineplot(data=spotify_data_csv, x='top year', y='pop', marker='o', ci=None)
plt.title('Average Popularity of Top 100 Songs Over the Years (2010-2019)')
plt.xlabel('Year')
plt.ylabel('Average Popularity')
plt.tight_layout()
plt.grid(True, which='both', linestyle='--', linewidth=0.5)
plt.show()
```



Summary:
The line chart depicts the average popularity of the Top 100 Spotify songs from 2010 to 2019. A prominent upward trend in song popularity is evident throughout this period, perhaps suggesting an expanding Spotify user base or shifting musical preferences. 2019 stands out as the zenith of average popularity, implying these tracks had a profound impact on Spotify's users.

4. Conclusion

The analysis provided offers a comprehensive exploration into a decade of musical evolution, as epitomized by Spotify's Top 100 songs from 2010 to 2019. By leveraging intricate datasets and utilizing advanced data visualization methodologies, this investigation reveals critical trends, genres, and song attributes that resonated profoundly with listeners. From the unparalleled supremacy of Dance Pop to the emotional spectrum captured by song valence, this study delves into the predilections of Spotify's vast global audience. Detailed histograms, line plots, and bar charts elucidate patterns and shifts in musical preferences, furnishing a wealth of insights. This analysis, appealing to both music enthusiasts and data analysts, offers a harmonious fusion of music and data, unveiling the captivating narrative behind the numbers.