Spotify Songs EDA

Problem Statement:

In the age of digital music streaming, platforms like Spotify have revolutionized the way we listen to music. Analyzing streaming data can reveal valuable insights into listening habits, genre popularity, song features, and trends that define the music industry. In this project, we will explore Spotify's track data to identify patterns, relationships between audio features and popularity, and the characteristics of top-charting songs.

Dataset Link:

https://www.kaggle.com/datasets/iamsumat/spotify-top-200-charts-2020

EDA Questions:

1 How many unique artists are in the dataset?

2 What is the distribution of song popularity scores?

3 Which artists appear most frequently in the Top 200?

4 What are the most common genres?

5 What is the average duration of songs?

6 How do audio features like danceability, energy, and valence vary across genres? 7

Which countries contribute the most to the top charts?

Visualization Questions:

1 Bar chart of Top 10 artists by number of chart appearances.

2 Histogram of song popularity scores.

3 Boxplot showing distribution of danceability across genres.

4 Heatmap showing correlation between audio features (danceability, energy, valence, tempo, loudness).

5 Line chart showing trend of popularity over time.

6 Scatter plot of energy vs loudness colored by genre.

7 Pie chart showing percentage distribution of songs by explicit vs non-explicit lyrics.