Project Title
Global Spotify Listener Trends

Team Name
007

Team Members

Anna Barbera, Tanner Horton, Jonathan Michel, Victor Pang, Bronwyn Milne

Project Description/Outline

Abstract: Analyze trends in Spotify Music data over time.

Objective: Make a Jupyter Notebook application that will group Spotify data by year, analyze trends and display them visually.

Question: Which markets hold the most power over global music trends?"

Hypothesis: The top streaming market will have the most substantial influence over global charts.

Null: The top streaming market will have the most substantial influence over global charts.

Alternative: The top streaming market will not have a substantial influence over global charts.

Final Four Questions:

1)Which countries were the most influential on the global Spotify market?

- Create dataframe with all of the countries and sort by total streams
- Plot findings in a bar graph
    - Calculate US streams over Global streams to find percent (Anna)

2)What are the top ranked songs (Bronwyn) and artists globally (Victor)?

3)How did the top streaming countries contribute to the most streamed song for each year?

4)What percentage of the top songs and artists in the top streaming country are a part of the global rankings?

Research Questions to Answer

1. Direction of genre trends based on top songs/artists (most popular genre per year)
2. Compare top songs by country
3. Compare top artists by country
4. Compare how individual songs performed globally
5. Which countries were the most influential on the global Spotify market?
6. Which songs were the most influential globally (broken down by year)?
7. Which genres were most influential on the global Spotify market (by year)?

8. Which years were most influential on the global Spotify market during this time?

Optional Questions

1. What genres typically appear in the top 200
2. How have the popular genres shifted in the top 200 over time
3. Which genres are popular in which countries
4. Most popular artists globally?
5. How do popular artists relate to top songs?
6. Has song duration changed over time?

CLEAN-UP TASKS:

1. Replace instances of ‘global’ with ‘regions’ excluding global and US
2. Add charts from Tanner and Anna

Presentation Order:

1. Anna - Introduction to EDA and brief overview of study relevancy
2. Jonathan - Which countries are the most influential w/results and graphs
3. Victor - Artists - US vs. global
4. Bronwyn -
5. Tanner - Songs - US vs. global and t-test?

Datasets to Be Used

https://www.kaggle.com/datasets/dhruvildave/spotify-charts?resource=download

https://www.kaggle.com/datasets/sashankpillai/spotify-top-200-charts-20202021

https://www.kaggle.com/datasets/paradisejoy/top-hits-spotify-from-20002019

https://www.kaggle.com/datasets/leonardopena/top-spotify-songs-from-20102019-by-year

https://www.kaggle.com/code/aeryan/spotify-music-analysi

https://www.kaggle.com/datasets/adnananam/spotify-artist-stats

Rough Breakdown of Tasks

Anna - GitHub Manager

Bronwyn - Presentation

Jonathan - Data Cleaning

Tanner - Data Manager

Victor - Data Visualization

Project Steps

1) Clean dataset
2) Create new CSV from clean data
3) Analyze music trends by time period for top songs
    a) Identify top performing songs by total number of streams
        i) (Remove global and sum all regions)
    b) Identify top performing artists by total number of streams
    c) Use Spotify API to determine each song's genre
        i) Identify top performing genres by total number of streams
    d) Create new dataframe with top findings
    e) Visualize top findings
4) Analyze music trends by region for top songs for each year (or a specific year)
    a) Identify top performing songs by total number of streams
    b) Identify top performing artists by total number of streams
    c) Use Spotify API to determine each song's genre
        i) Identify top performing genres by total number of streams
    d) Create new dataframe with top findings
    e) Display top findings on a map