Spotify: What Makes a Top Song?

DigHum 100 Dr. Anderson | Gloria Yang | May 27, 2021

Description: I've noticed that features of the most popular songs have changed a lot over the decades. While there are influences and consistencies in popular song attributes from the 20th century, the 21st century brought up new genres and styles of music. Modernization has led to different approaches and greater diversity of sounds such as louder sounds, faster tempos, and speechiness. In this project, I will be analyzing the popular songs on Spotify over time and looking for common attributes and trends.

Dataset: "Spotify: All Time Top 2000 Mega Dataset" is a dataset taken from Kaggle which contains audio statistics and attributes for Spotify's top songs. The dataset is contained within a csv file and includes data on 1995 songs released from 1959 to 2019. The data set breaks down each song by the following attributes: genre, year, BPM, energy, danceability, loudness, liveness, valence, length, acousticness, speechiness, and popularity.

I will be able to cluster and manipulate the dataset's columns to make our visualizations more clear during exploration. I will be using the dataset in order to explore the attributes of Spotify's most popular songs and the changes in these attributes over time.

Index	Title	Artist	Top Genre	Year	Minute (BPM)	Energy	Danceability	Loudness (dB)	Liveness	Valence	Length (Duration)	Acousticness	Speechiness	Popularity
1	Sunrise	Norah Jones	adult standards	2004	157	30	53	-14	11	68	201	94	3	71
2	Black Night	Deep Purple	album rock	2000	135	79	50	-11	17	81	207	17	7	39
3	Clint Eastwood	Gorillaz	alternative hip hop	2001	168	69	66	-9	7	52	341	2	17	69
4	The Pretender	Foo Fighters	alternative metal	2007	173	96	43	-4	3	37	269	0	4	76
5	Waitin' On A Sunny Day	Bruce Springsteen	classic rock	2002	106	82	58	-5	10	87	256	1	3	59
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1990	Heartbreak Hotel	Elvis Presley	adult standards	1958	94	21	70	-12	11	72	128	84	7	63
1991	Hound Dog	Elvis Presley	adult standards	1958	175	76	36	-8	76	95	136	73	6	69
1992	Johnny B. Goode	Chuck Berry	blues rock	1959	168	80	53	-9	31	97	162	74	7	74
1993	Take Five	The Dave Brubeck Quartet	bebop	1959	174	26	45	-13	7	60	324	54	4	65
1994	Blueberry Hill	Fats Domino	adult standards	1959	133	50	49	-10	16	83	148	74	3	56

Hyperlink to Colab:https://colab.research.google.com/drive/1DsZ4s_TyKT_Tudc2z3GnCYQGoZT_HXB?usp=
Hyperlink to CSV: https://drive.google.com/file/d/16BOmSQKc0WRAMaaHU9TBXTGERMbiz2uI/view?usp=sharing

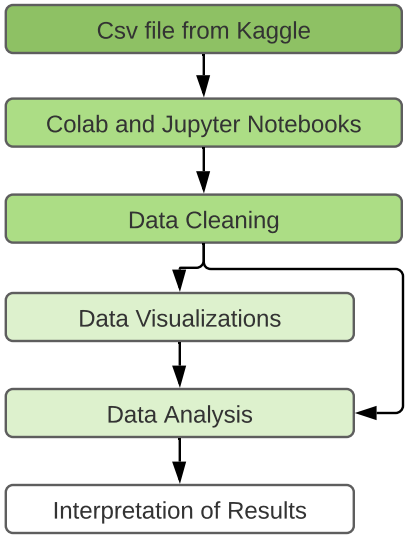
Questions for EDA:

- Can we predict a song's popularity based on attributes such as danceability, valence, loudness, and length?
- Are there any trends in the popularity of different music genres?
- How do song attributes differ according to different genres?
- How have the attributes of songs that appear on Spotify's Top 2000 List changed over time?

Tools:

- **Matplotlib and Seaborn:** Tool to create data visualizations on Python
- **Pandas:** Manipulating Spotify data in dataframes
- **Colab and Jupyter Notebooks:** Tool for data cleaning, data manipulation, and developing visualizations

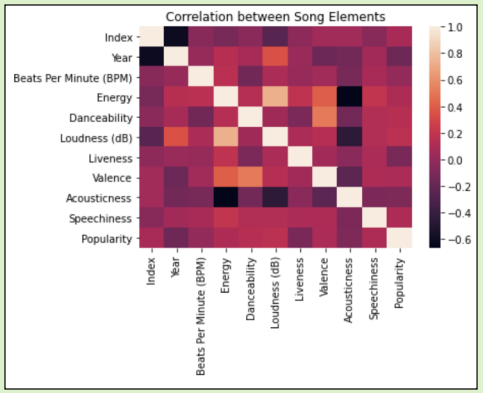
Methods: The plan is to create visualizations within the data from the questions asked earlier using Colab. I will be exploring correlations and trends in relation to song attributes including BPM, loudness, and danceability. With the feedback of my classmates and Dr. Anderson, I will continue to iterate upon my data analysis.



Interpretation of Results:

At the end of my data analysis, I expect to find that the genre of Spotify's top songs will have shifted to genres like hip hop, rap,and pop. Due to my research on the shift of music trends over the 20th to 21st century, I also expect to see shifts in certain attributes such as loudness, liveness, and energy. I do however anticipate that there will be consistencies in attributes like danceability that will allow us to predict the likelihood of another song's popularity.

I've started playing around with my dataset to get a beginning idea of existing correlations that I can explore regarding song attributes; a heatmap that I created is shown on the right!



Works Cited

https://www.kaggle.com/iamsumat/spotify-top-2000s-mega-dataset
https://sites.google.com/site/musicpopularitydataanalysis/trends-over-time
https://towardsdatascience.com/predicting-spotify-song-popularity-49d000f254c7
https://techxplore.com/news/2019-09-spotify-songs.html