

Topic: Music

Question: What were the top most played songs/genres during the peak of the COVID-19 Pandemic (2020-2021)?

Aim: Since music is basically a form of therapy or escape across generations, this data collection aims to survey what kind of music or songs were used to cope with the stress of the unforeseen COVID-19 Pandemic.

Data Collection Plan

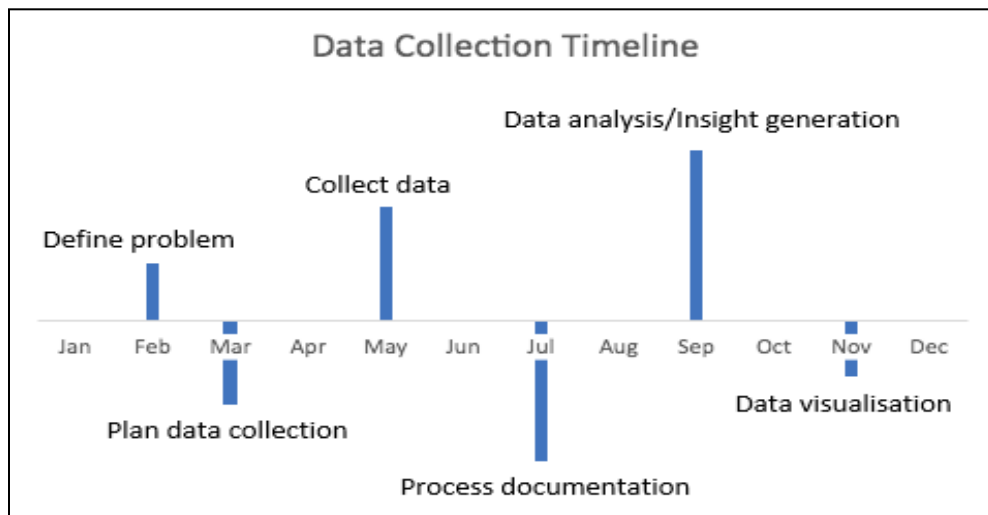
Type of data needed:

- Quantitative, able to measure the number of streams a song/genre has.
- Number of streams will indicate its popularity among listeners.
- Why use secondary data:
 - Save time.
 - Minimise cost and effort.
 - Convenience, since music is also generally accessible and available to the public.

Data collection method:

- Web scraping
- Allows quick and efficient extraction of data.
- Data Source:
 - Kaggle
 - A good platform to start and learn for beginners
 - Numerous datasets available can be found for exploration
- Tools:
 - Google Sheets
 - Excel Spreadsheets
 - Kaggle

Timeline chart



Metadata

The dataset contains information on all the songs that have been on the Top 200 Weekly (Global) charts of Spotify in 2020 & 2021. The data collection method was web scraping. Information about the song title and artist was retrieved from spotifycharts.com and then scraped using BeautifulSoup. Later, it was then given to the Spotipy Python library to get more information like the audio features. Description of data is included.

The dataset was prepared by a secondary source, Sashank Pillai in 2021.

Column	Data type	Description
Song Name	String	Name of the song
Streams	Integer	Number of times the song has been played
Highest Charting Position	Integer	The highest position that the song has been on in the Spotify Top 200 Weekly Global Charts in 2020 & 2021.
Number of Times Charted	Integer	The number of times that the song has been on in the Spotify Top 200 Weekly Global Charts in 2020 & 2021.
Artist	String	The main artist of the song.

Artist Followers	Integer	The number of followers the artist has on Spotify.
Genre	String	The genres the song belongs to.
Release Date	Date	The initial date that the song was released.
Popularity	Integer	The popularity of the track. The value will be between 0 and 100, with 100 being the most popular.
Danceability	Float	Describes how suitable a track is for dancing. A value of 0.0 is the least danceable and 1.0 is the most danceable.
Energy	Float	Measures from 0.0 to 1.0 and represents a perceptual measure of intensity and activity.
Loudness	Float	The overall loudness of a track in decibels (dB).
Speechiness	Float	Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording, the closer to 1.0 the attribute value.
Acousticness	Float	A measure from 0.0 to 1.0 whether the track is acoustic.
Liveness	Float	Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live.
Tempo	Float	The overall estimated tempo of a track in beats per minute (BPM).
Valence	Float	A measure of 0.0 to 1.0, describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive.

Reference

Pillai, S. (2021). *Spotify Top 200 Charts (2020-2021)* [Data set]. Kaggle. <https://doi.org/10.34740/KAGGLE/DSV/2529719>

License

Data files © Original Authors